

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```
Run on:      Fri Sep 3 13:11:46 1999;  MasPar time 18.55 Seconds
```

Title: >US-09-120-044-3-COPY

```

Description: (1-4/1) From devil20044.pcp
Perfect Score: 3234

```

1 MANKAVNDFILAMNYDKKL.....TISWGTLLYPQVEDKVEND 471

Gap 11

```
post-processing: Minimum Match 0%
                  Listing first 45 summaries
```

a-geneseq35

8:part8 9: part2 9:part5 9:part5 5:part5 6:part6 7:part7
14:part14 10:part10 10:part10 11:part11 12:part12 13:part13
15:part15 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result

NO.	Score	Match	Length	DB	ID	Description	Pred.	No
1	3332	99.9	471	1	R05924	Immunogenic pneumolys	6.18e-27	1
2	3056	94.5	411	1	R05924	Immunogenic pneumolys	3.57e-25	2
3	1416	43.8	480	6	R33842	Streptolysin O variant	5.45e-11	3
4	1398	43.2	480	6	R33841	Soluble streptolysin	2.13e-10	4
5	1391	43.0	571	2	R10376	Streptolysin O deriva	8.88e-10	5
6	1097	33.9	571	2	R06090	Bacteriophage lambda	7.61e-83	6
7	367	11.3	113	33	W60952	Streptococcus pneumoniae	7.90e-20	7
8	108	3.5	117	17	R88185	Listeriolysin O haemo	5.96e-09	8
9	112	3.3	299	33	W60976	Streptococcus pneumoniae	1.10e-01	9
10	107	3.3	1228	25	W22862	Bacillus stearothermo	1.28e-01	10
11	107	3.3	1228	17	R77673	S-layer protein encod	1.38e-01	11
12	101	3.1	188	29	W55347	H. pylori ORF 01ce116	3.16e-01	12
13	101	3.1	148	29	W55347	H. pylori ORF 29gc303	3.16e-01	13
14	101	3.1	455	28	W20753	H. pylori transpore	3.16e-01	14
15	96	3.0	667	8	R1781	Rabbit HBI protein wh	6.61e-01	15
16	94	2.9	114	2	R07691	Myobacterial 65kd ant	8.85e-01	16

17	93	2.9	459.20	W04722	Aromatic acyl transferase	1.02e+02
18	94	2.9	555.33	W30676	Glycerol dehydratase	8.85e+01
19	95	2.9	555.32	W62064	Klebsiella pneumoniae	8.85e+01
20	96	2.9	649.38	W81977	Ellichia sp. B3 prot	7.65e+01
21	95	2.9	906.35	W12290	Potato starch branchi	7.65e+01
22	95	2.9	906.35	W69300	Potato starch branchi	7.65e+01
23	94	2.9	1057.36	W72067	Potato class B starch	7.65e+01
24	94	2.9	1114.36	W72205	HSV-2 strain S85 cont	8.85e+01
25	94	2.9	1114.36	W72096	HSV-2 strain S85 cont	8.85e+01
26	93	2.9	1248.34	W19783	HSV-2 strain S85 cont	8.85e+01
27	90	2.8	210.15	R180623	Human multiple regu	1.02e+02
28	91	2.8	226.20	W1582	Coho Salmon growth ho	1.57e+02
29	90	2.8	355.12	R64226	Streptomyces pristina	1.36e+02
30	90	2.8	493.3	R13992	Human 38kDa FK-506 bi	1.36e+02
31	92	2.8	527.2	R05079	P. falciparum sporozo	1.57e+02
32	92	2.8	527.33	W61278	Human monomamine oxida	1.57e+02
33	92	2.8	556.17	R96737	Monamine oxidase A.	1.18e+02
34	90	2.8	625.23	W18008	A. niger Bo-1 carboxy	1.18e+02
35	89	2.8	625.23	W19490	Arabidopsis CEB1 prot	1.57e+02
36	90	2.8	713.3	R13991	App695 mutant A-beta-	1.81e+02
37	89	2.8	751.23	W19505	Plasmodium falciparum	1.57e+02
38	89	2.8	751.23	W19452	App751 mutant A-beta-	1.81e+02
39	89	2.8	770.23	W19300	App751 mutant A-beta-	1.81e+02
40	89	2.8	1066.17	R91300	APP770 mutant A-beta-	1.81e+02
41	91	2.8	1093.8	R42818	TRPA box blinding prot	1.81e+02
42	92	2.8	1244.24	W19785	TME.	1.36e+02
43	92	2.8	126.24	W19784	Mouse multiple regu	1.18e+02
44	89	2.8	186.23	W24790	Human multiple regu	1.18e+02
45	91	2.8	1810.16	R45563	P. falciparum liver s	1.81e+02
46	91	2.8	1810.16	R45563	Chicken erythrocyti	1.81e+02

TEST	RESULT	TEST FOR
1		

AC R05923; 4/1 AA.

DT 23-NOV-1990 1411st out...

Immunogenic pneumolysin variant.

Streptococcus pneumoniae; meningitis; bacteremia; ds

28-JUN-1990

PE 15-DEC-1989; AU0539.

PA (PATO/) PATON J C.

11 FASHION CC, HANSMAN DJ, BOULNOIS GJ, ANDREW PW, MITCHELL TJ
12 WALKER JA.
13

WPI; 90-224494/29

PT New non-toxic,

protective vaccines against *Streptococcus pneumoniae*, and DNA sequences encoding the

Claim 5; Page 11; 25pp; English

making them useful in vaccination against pneumolysin,

meningitis, bacter

Best Local Similarity	97.98%	Score	3232	DB	1	Length	471
Best Local Similarity	97.98%	Prod	5	10	273		

Matches	461;	Conservative	1;	Mismatches	9;	Indels	0;	Gaps	0
---------	------	--------------	----	------------	----	--------	----	------	---

1 mankavndfilamnydkkkllthqeslenrfikeanqjndefvvarlvknlstataa: ca

1 MANKAVNDETI,AMNYDKRKI,THOCSEIENE-E-VTCOR DRTT... IT.....

61 2014-01-01 00:00:00

[illegible]

FALL VDEI LEXENN PILLAV DRAPMTYSIDLPGGLASSDSFLQVEDPSN 120

121 ssvrgavndllakwhgdgyqvnvparmqekitahsmeglkvkfqsdfecktansldidf 180

121 SSVRGAVNDLAKWHODYGOVNNVPARXOYEKTAHSMEOIKYKESDEFEKTCVCTDTT

Db 181 nsvhsgekqigivnfkqilytvsdvavknpgdfqdtvtedlkgqisaeerplyissv 240
 |||||
 Qy 181 NSVHSEKQIOIVNKKOIIYTVSDAVKNPGDFQDTVTEDELKQRGISAEERPLYISXV 240
 Db 241 aygrqyylkettsksdeveaafaelikgvkvaoptekqilndtevkavllggdpsssa 300
 |||||
 Qy 241 AYGRQYVILKETTSKSXEVEAFAELIKGVKVAOPTEKQILNDTEVKAVLLGGDPSSGA 300
 Db 301 rvtgkvdmvedlqgesrftadhpglpisyttsfjrdnvavafqnsctdyetkvtayrn 360
 |||||
 Qy 301 RVTGKVDMEDELQGESRFTADHPGLPISYTTSELRDNVAVAFQNSDYETKVTAYRN 360
 Db 361 gdlldhsgayvaayyitfdelsfhnbgkevltpkafdnngdltahftsipkgnvyrn 420
 |||||
 Qy 361 GDLLDHSGAYVAOYYITWNELSYDHQKEVLTPKAMDNGODLTAHFTTSLPKGNVRN 420
 Db 421 lsvkirectglawewrtyektdplvrkrttslqwtllypvedkvead 471
 |||||
 Qy 421 LSVKIRECTGLAWEMWRTYKTDPLVRKRRTISLQWTLTYPVEDKVEAD 471

RESULT 2

ID R05924 standard; protein: 471 AA.
 AC R05924:
 DT 23-NOV-1990 (first entry)
 DE Immunogenic pneumolysin variant.
 KW Streptococcus pneumoniae.
 OS Streptococcus pneumoniae.
 FH Key location/Qualifiers
 FT misc_difference 428
 FT label=G, A, S
 FT misc_difference 435
 FT label=Q, D
 PN W0506951-A.
 PD 28-JUN-1990.
 PE 15-DEC-1989; AU0539.
 PR 16-DEC-1989; AU-001989.
 PA (PATON/J. PATON J. C.
 PI PATON JC, HANSMAN DJ, BOULNOIS GJ, ANDREW PW, MITCHELL TJ,
 PI WALKER JA.
 DR WPI; 90-224494/29.
 PT New non-toxic, immunogenic mutants of pneumolysin - useful in
 PT protective vaccines against Streptococcus pneumoniae, and DNA
 PT sequences encoding them.
 PS Disclosures: P: English.
 CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
 CC making them useful in vaccination against pneumonia and associated
 CC meningitis, bacteraemia etc.
 SQ Sequence 471 AA:

Query Match 94.5%; Score 3056; DB 1; Length 471;

Best Local Similarity 94.5%; Pred. No. 3,57e-257;

Matches 445; Conservative 6; Mismatches 20; Indels 0; Gaps 0;

Db 1 mankavndfilamnydkkllthggesienrfikegnqlpdefvixekrrkststsd 60
 |||||
 Qy 1 MANKAVNDFILAMNYDKKLLTHGGESIENRFKEGNQLPDEFVIXEKRRKSLSTSDI 60
 Db 61 svratdstrlygalllvdeetlennptllavdrpmtyssldlpjassdsflgvedpsn 120
 |||||
 Qy 61 XVRATDSRLYGPALLVDEETLXENNPILLAVDRAPMTYSIDLPGCLASDSDFLQVEDPSN 120
 Db 121 ssyrgavnnllakwbhdygqvnvparnmyekktahsmeglkvfkgsdfektgnsltdf 180
 |||||
 Qy 121 SSRGAVNDLLAKWHDYGOVNNVPAKQYKXITAHSMEOULKVKGSDFEKTSGLDIDF 180
 Db 181 nsvhsgekqigivnfkqilytvsdvavknpgdfqdtvtedlkgqisaeerplyissv 240
 |||||
 Qy 181 NSVHSEKQIOIVNKKOIIYTVSDAVKNPGDFQDTVTEDELKQRGISAEERPLYISXV 240
 Db 241 aygrqyylkettsksdeveaafaelikgvkvaoptekqilndtevkavllggdpsssa 300
 |||||

Qy 241 AYGRQYVILKETTSKSXEVEAFAELIKGVKVAOPTEKQILNDTEVKAVLLGGDPSSGA 300
 Db 301 rvtgkvdmvedlqgesrftadhpglpisyttsfjrdnvavafqnsctdyetkvtayrn 360
 |||||
 Qy 301 RVTGKVDMEDELQGESRFTADHPGLPISYTTSELRDNVAVAFQNSDYETKVTAYRN 360
 Db 361 gdlldhsgayvaayyitfdelsfhnbgkevltpkafdnngdltahftsipkgnvyrn 420
 |||||
 Qy 361 GDLLDHSGAYVAOYYITWNELSYDHQKEVLTPKAMDNGODLTAHFTTSLPKGNVRN 420
 Db 421 lsvkirectglawewrtyektdplvrkrttslqwtllypvedkvead 471
 |||||
 Qy 421 LSVKIRECTGLAWEMWRTYKTDPLVRKRRTISLQWTLTYPVEDKVEAD 471

RESULT 3

ID R33842 standard; protein: 480 AA.
 AC R33842:
 DT 15-JUL-1993 (first entry)
 DE Streptolysin O variant msLO.3/6.
 KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
 KW Streptococcus pyogenes; ASO.
 OS Streptococcus pyogenes.
 PN W0930515-A.
 PD 18-MAR-1993.
 PE 03-AUG-1992; U06380.
 PR 30-AUG-1991; US-752428.
 PA (BECK) BECKMAN INSTR INC.
 PI Adams CW;
 PI WPI; 93-100979/12.
 DR N-PSDB: Q38287.
 PT Streptolysin O variants produced by recombinant DNA technology -
 PT having no haemolytic activity and recognised by wild-type
 PT anti-streptolysin O antibodies, useful in diagnosis of
 PT Streptococcus pyogenes infection
 PS Claim 17; Fig 2; 80pp; English.
 CC The sequence represents a soluble variant of Streptolysin O (SLO)
 CC having no haemolytic activity. The variant SLO can specifically
 CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
 CC activity. It can be used for the determination of previous and/or
 CC current infection by Streptococcus pyogenes. It enables ASO assays
 CC to become commercially viable and avoids the risk of handling haemo-
 CC lytically active S. pyogenes wild-type SLO.
 SQ Sequence 480 AA:

Query Match 43.8%; Score 1416; DB 6; Length 480;

Best Local Similarity 42.0%; Pred. No. 5,45e-111;

Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

Db 18 indkyslnpnelevlakngelienfvpkegyvkkadkfivierkkknltprvdisids 77
 |||||
 Qy 6 VNDFILAMNYDKKLLTHGGESIENRFKEGNQLPDEFVIXEKRRKSLSTSDIXVAT 65
 Db 78 vtdtlypaalqankgtentkpkdavytknpgkibldpgmddkt-vevndptyanst 136
 |||||
 Qy 66 NDSRLYGPALLVDEETLXENNPILLAVDRAPMTYSIDLPGCLASDSDFLQVEDPSN 125
 Db 137 aichlvngwhnysggnltlparqytksmysksqiaaelnynskllldtltgldfksisk 196
 |||||
 Qy 126 AVNDILAKWHDYGOVNNVPAKQYKXITAHSMEOULKVKGSDFEKTSGLDIDFNSVHS 185
 Db 197 gekvmlaaykqifvytsaanlpnradvfdksvtfkelqtkyvsneaprlfyfnyvayrt 256
 |||||
 Qy 166 GEKOIOIVNKKOIIYTVSDAVKNPGDFQDTVTEDELKQRGISAEERPLYISXVAYGRQ 245
 Db 257 vfvketesksendeaafaelikgvkvaoptekqilndtevkavllggdpsssa 316
 |||||
 Qy 246 VYKLETTSSXEVEAFAELIKGVKVAOPTEKQILNDTEVKAVLLGGDPSSGARVYVG 305
 Db 317 dfcdvlnvkdnatfstrknepayisytvflknkklagvnnarteyetstetysgkln 376
 |||||
 Qy 306 KYDWMEDLQGESRFTADHPGLPISYTTSELRDNVAVAFQNSDYETKVTAYRNGLDLL 365

Db 377 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 436
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTHFTTSLPLKGNVNLVSKI 425
 Db 437 rectglawewrkylderdvklskelnvngstlsp 473
 QY 426 RECTGLAWEMWRKYEKTDLPVLRKRRTISWGTTLTP 462

RESULT 4
 ID R33841 standard; Protein; 480 AA.
 AC R33841;
 DT 08-JUL-1993 (first entry)
 DE Soluble streptolysin O variant.
 KW Recombinant; haemolytic activity; immunodiagnostic activity; ss.
 OS Streptococcus pyogenes.
 PN W09305156-A.
 PD 18-MAR-1993.
 PF 31-JUL-1992; 006398.
 PR 30-AUG-1991; US-752429.
 PA (BECI) BECKMAN INSTR INC.
 PI Adams CW, Wang EY;
 DR WPI: 93-100980/12.
 DR N-PSDB; 038286.
 PT Deriv. of streptolysin O with haemolytic activity - used in
 PT immuno-diagnostic assays which rely on haemolytic activity of
 PS wild-type Streptolysin O
 PS Claim 7; Fig 2; 60pp; English.
 CC The sequence is that of a soluble deriv. of Streptolysin O
 CC with haemolytic properties which was obtd. from Streptococcus
 CC genomic library clone r50.3. It is useful in immunodiagnostic assays
 CC which rely upon, e.g. the haemolytic properties of wild type
 CC Streptolysin O. Recombinant streptolysin O is obtd. more cheaply
 CC than purified streptolysin O obtd. from Streptococcus pyogenes.
 SQ Sequence 480 AA;

Query Match 43.28; Score 1398; DB 6; Length 480;
 Best Local Similarity 41.6%; Pred. No. 2.13e-109;

Matches 190; Conservative 104; Mismatches 162; Indels 1; Gaps 1;

Db 18 indkylslnynelevlakngetienvfpekgvkkadkfivierkkknintpvdslsids 77
 QY 6 VNDFLIAMYDKKKLLTHOGESIEENRFXKEGNQLPDEFVYXERKKRSLSINTSDIXVTAT 65
 Db 78 vdrtypaalqlankgftenkpdavvtkrnpqknhidlpmgdkat-vevndptyanvt 136
 QY 66 NDSRLYPGALLVYDETLXENNPFTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSVVG 125
 Db 137 aldnlvngwhndysggnltpartqytesmvsksqieaalnvnskildgtlgidfksisk 196
 QY 126 AVNDLLAKWHQDYGQVNNVPAKQYERKTAHSMEOIKVFCGSDFEKTGNSLIDFNSVHS 185
 Db 197 gekkvmiaaykqgifytvsanlpnpadvdksvtfkelgrkysneapplifvsnaaygrt 256
 QY 186 GEKOIOIVNKKOIIYTVSDAVAKNPGDVFQDTYVEDLKQKGISMERLVYISXAAYGRQ 245
 Db 257 vfvkletsksndveaafsaalkgtdvktngkysdliensftavvlggdaaenhkvvtk 316
 QY 246 VYKLETTSSXSEVEAFAELIKGVKAPQTEWKQILDTEVKAAYILGGDPSSGARVYTG 305
 Db 317 dfdvtrnvikdnatfsrknpayisvtsvflknkknagvnmrtteyvetsteysgknl 376
 QY 306 KVDWEDLIOEGSRFTADHPGLPISTTSFLRDNVAVTFQNSTDYVEKTVAYRNGLDLL 365
 Db 377 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 436
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTHFTTSLPLKGNVNLVSKI 425
 Db 437 rectglawewrkylderdvklskelnvngstlsp 473
 QY 426 RECTGLAWEMWRKYEKTDLPVLRKRRTISWGTTLTP 462

RESULT 5
 ID R10376 standard; Protein; 571 AA.
 AC R10376;
 DT 05-APR-1991 (first entry)
 DE Streptolysin O derivative.
 KW SLO.
 OS Streptococcus pyogenes.
 PN GB2233977-A.
 PD 23-JAN-1991.
 PF 04-JAN-1989; 000107.
 PR 04-JAN-1989; GB-000107.
 PA (KERO) KEROE M.
 PI (PINK) PINKNEY M.
 PI Kehoe M, Pinkney M;
 DR WPI: 91-024598/04.
 DR N-PSDB; 010320.

PT Deriv. of thiol-activated protein streptolysin O - contg. no
 PT cysteine amino acid but retaining cytolitic activity, used for
 PT detecting antibodies in samples
 PS Disclosure; Fig 2; 9pp; English.
 CC The SLO derivative contains no cysteine residues, with cytolitic
 CC activity giving resistance to inactivation by oxidation or
 CC thiol groups. Abs raised to the SLO may be used in the decton and
 CC diagnosis of Streptococcus pyogenes infection.
 SQ Sequence 571 AA;

Query Match 43.0%; Score 1391; DB 2; Length 571;
 Best Local Similarity 41.6%; Pred. No. 8.88e-109;

Matches 190; Conservative 104; Mismatches 162; Indels 1; Gaps 1;

Db 109 indkylslnynelevlakngetienvfpekgvkkadkfivierkkknintpvdslsids 168
 QY 6 VNDFLIAMYDKKKLLTHOGESIEENRFXKEGNQLPDEFVYXERKKRSLSINTSDIXVTAT 65
 Db 169 vdrtypaalqlankgftenkpdavvtkrnpqknhidlpmgdkat-vevndptyanvt 227
 QY 66 NDSRLYPGALLVYDETLXENNPFTLLAVDRAPMTYSIDLPGLASSDSFLQVEDPSNSVVG 125
 Db 228 aldnlvngwhndysggnltpartqytesmvsksqieaalnvnskildgtlgidfksisk 287
 QY 126 AVNDLLAKWHQDYGQVNNVPAKQYERKTAHSMEOIKVFCGSDFEKTGNSLIDFNSVHS 185
 Db 288 gekkvmiaaykqgifytvsanlpnpadvdksvtfkelgrkysneapplifvsnaaygrt 347
 QY 186 GEKOIOIVNKKOIIYTVSDAVAKNPGDVFQDTYVEDLKQKGISMERLVYISXAAYGRQ 245
 Db 348 vfvkletsksndveaafsaalkgtdvktngkysdliensftavvlggdaaenhkvvtk 407
 QY 246 VYKLETTSSXSEVEAFAELIKGVKAPQTEWKQILDTEVKAAYILGGDPSSGARVYTG 305
 Db 408 dfdvtrnvikdnatfsrknpayisvtsvflknkknagvnmrtteyvetsteysgknl 467
 QY 306 KVDWEDLIOEGSRFTADHPGLPISTTSFLRDNVAVTFQNSTDYVEKTVAYRNGLDLL 365
 Db 468 shgagayvaqyelwdeinyddkqkveitkrrwdnmysktsfvtvlpjgansnrima 527
 QY 366 DHSGAYVAQYIITWNLSTYDHQGEKVLTPKAWDRNGODLTHFTTSLPLKGNVNLVSKI 425
 Db 528 rectglawewrkylderdvklskelnvngstlsp 564
 QY 426 RECTGLAWEMWRKYEKTDLPVLRKRRTISWGTTLTP 462

RESULT 6
 ID R06000 standard; Protein; 371 AA.
 AC R06000;
 DT 22-NOV-1990 (first entry)
 DE Bacteriophage lambda PL promoter - streptolysin O (SLO) fusion
 DE protein of plasmid pmk306.
 KW Streptolysin O; SLO; protease; bacteriophage lambda; ds.
 OS Streptococcus pyogenes.
 FT Key Location/Qualifiers
 cleavage_site 33..34

RESULT 9
ID M60976 standard; Protein; 299 AA.
AC M60976:
DE 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae encoded polypeptide.
KW coding region; ORF; open reading frame; antibacterial;
KW infection; prevention; meningitis.
OS Streptococcus pneumoniae.
PN M09819689-A1.
PD 14-MAY-1998.
PF 27-OCT-1997; 019226.
PR 01-NOV-1996; US-029930.
PA (SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PI Black MT, Hodson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PW;
PI MPI: 98-286586/25.
DR N-PSDB; V37371.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PS Claim 11: Page 71-72; 130pp; English.
CC The sequence is that of the polypeptide encoded by a region isolated
CC from S. pneumoniae which shows homology to nicotinate-nucleotide pyro-
CC phosphorylase. It, or agonists of it, may be useful as an anti-
CC bacterial for treatment or prevention of infection, specifically caused
CC by S.pneumoniae (particularly meningitis) but possibly also Helicobacter
CC pylori (ulcers and gastric cancer). It may be of particular
CC use before insertion of an in-dwelling device or any other
CC invasive procedure. The protein, or nucleic acid encoding
CC it, can also be used in vaccines to induce a cellular
CC and/or humoral immune response, or to screen for other
CC antibacterials. The DNA may also contain flanking sequences
CC that are potential sources of control elements for bacterial
CC gene expression. Detecting a sequence encoding the protein
CC can be used diagnostically, e.g. to detect a mutation for
CC serotyping or classifying infectious agents.
SQ Sequence 299 AA;

Query Match 3.3%; Score 108; DB 33; Length 299;
Best Local Similarity 21.4%; Pred. No. 1.10e+01;
Matches 22; Conservative 35; Mismatches 40; Indels 6; Gaps 6;

Db 38 dystalafdhgqakvsjfkagaylag-lwfgvflfaeyrfqphgfdgdrllts 96
QY 14 NYDKRKLTLHOGESIENRFXKGNQLPDEFVYXERKRSLSLTNS-DIXVATATDSRLYP 72
DB 97 gdl-vl-elig-svrslltcevalnfqlhslgiasmtaayve 136
QY 73 GALLVVDLTLXENNPFTLLAVDRAPATYISIDPLGLAS-SDSFLQ 114

RESULT 10
ID W22862 standard; Protein; 1228 AA.
AC W22862:
DE 27-FEB-1998 (first entry)
DE Bacillus stearothermophilus S-layer protein sbs-A.
KW S-layer; sbs-A; vaccine; adjuvant; carrier; hybridisation assay;
KW molecular spinning nozzle; molecular laser.
OS Bacillus stearothermophilus.
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= sig_peptide
FT Peptide 31..1228
FT /label= mat_peptide
PN DE19603649-A1.
PD 07-AUG-1997.
PF 01-FEB-1996; 003649.
PR 01-FEB-1996; DE-003649.
PA (LUBITZ) LUBITZ W.
PA (SLEY) SLEYTR U.
PI Kuen B, Lubitz W, Sleytr U;

DR MPI: 97-394558/37.
DR N-PSDB; T75487.
PT Preparation of S-layer proteins by expressing sbs-A gene in Gram
PT negative bacterium or new sbs-B gene in any host, also new
PT recombinant proteins containing heterologous inserts, e.g.
PT epitope(s), useful as vaccines and adjuvants
PS Claim 1; Pages 9-14; 31pp; German.
CC The present sequence is the Bacillus stearothermophilus pV72
CC S-layer protein, sbs-A. S-layer structures can be used as vaccines
CC or adjuvants, particularly when they include a bacterial ghost that
CC may contain additional epitopes in its membrane. Other uses of
CC recombinant sbs-A, depending on the nature of the inserted peptide,
CC are as an universal carrier for biotinylated reactants for use in
CC immunological or hybridisation assays (the insert is streptavidin),
CC to induce immune responses (epitopes), as a reagent for removing
CC cytokine or toxin from serum (antigenic epitopes), as a molecular
CC spinning nozzle (polyhydroxybutyrate synthase) and as a molecular
CC laser (luciferase).
SQ Sequence 1228 AA;

Query Match 3.3%; Score 107; DB 25; Length 1228;
Best Local Similarity 23.5%; Pred. No. 1.28e+01;
Matches 39; Conservative 44; Mismatches 73; Indels 10; Gaps 10;

Db 200 ydliv-amkar-ev-qdgvkagnldkakaavdginyjpkvtdafteit-evakkalda 255
QY 200 YTVSDAVKKNPQVDFQDTVEDLKORGISAERPLVYISXVARGVYKLERTSXSKEY 259
DB 256 deaalfpkvesvs-ahntqkavelcavpntg-lklqisaanedtvnvtvrylkvdg 313
QY 260 -EAAFEALKGVAVAPQTEWKQI-LDNTKRAVILLAGDPSGGRVVTGGRVWEDLIORG 317
DB 314 nlfafntadvstldgklitvdastpfennetkykvvgixdkng 359
QY 318 S-RETDHPGLPISYTTSTFLRDVNVATFQNSIDY-VETVATVAVRNG 361

RESULT 11
ID R77673 standard; Protein; 1228 AA.
AC R77673:
DE 16-JUL-1996 (first entry)
DE S-layer protein encoded by sbsA gene.
KW sbsA gene; S-layer; cell surface layer; expression.
OS Bacillus stearothermophilus.
FH Key Location/Qualifiers
FT Peptide 1..30
FT /label= sig_peptide
FT Protein 31..1228
FT /label= mat_protein
PN DE4425527-A1.
PD 25-JAN-1996.
PF 19-JUL-1994; 425527.
PR 19-JUL-1994; DE-425527.
PA (VOGE-) VOGELBUSCH GMBH.
PI Lubitz W;
PI MPI: 96-077933/09.
DR N-PSDB; T08695.
PT Nucleic acid encoding signal peptide of Bacillus stearothermophilus
PT S-layer protein - which has a lysine content of at least 10 per
PT cent.
PS Disclosure: Page 6-7; 12pp; German.
CC The claimed signal peptide encoding sequence (see tag e of T08695)
CC is pref. operably linked at the 3' end to a protein encoding
CC sequence. The protein is pref. the S-layer protein (see tag f of
CC T08695). At the 5' end, the signal peptide encoding sequence is
CC pref. linked to an expression control sequence, pref. the sequence
CC given in T08696.
SQ Sequence 1228 AA;

Query Match 3.3%; Score 107; DB 17; Length 1228;
Best Local Similarity 23.5%; Pred. No. 1.28e+01;
Matches 39; Conservative 44; Mismatches 73; Indels 10; Gaps 10;

PR 01-APR-1996; US-630405.
 PA (ASTR) ASTRA AB.
 PI Berglindh OT, Smith D, Mellgaard BL.
 DR WPI: 97-052306/05.
 DR N-PSDB: T68006.
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PS Claim 56; Page 1166-1167; 1481pp: English.
 CC The present sequence is a Helicobacter pylori transporter protein.
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 SO Sequence 667 AA;

Query Match 3.1%; Score 101; DB 22; Length 667;
 Best Local Similarity 26.9%; Pred. No. 3,16e+01;
 Matches 14; Conservative 15; Mismatches 22; Indels 1; Gaps 1;

DB 533 lkgldivgvgfkgvgrantsda-aglmgletylelkpqnwkeklxykev 583
 Oy 237 ISXVAYGROYVLTLETTTSXSEVEAFALIKGVAVAPQTEMKOILDWTEVK 288

RESULT 15
 ID R41781 standard; Protein; 458 AA.
 AC R41781;
 DE 28-MAR-1994 (first entry)
 DE Rabbit HBI protein which binds hsp90.
 KW steroid hormone receptor; immunosuppressor; FK506; binding protein;
 KW FKBP; heat shock protein; hsp90; chaperone protein;
 OS rotamase activity; ss.
 OS Oryctolagus cuniculus.
 FH Key
 FT Location/Qualifiers
 FT domain
 FT 32..138
 FT /label= HBI-I
 FT /note= "immunophilin domain; has 49 per cent
 FT identity with FKBP without introducing
 FT insertions or deletions into the sequence"
 FT 139..148
 FT /label= hinge
 FT 149..253
 FT /label= HBI-II
 FT /note= "immunophilin domain; has 28 per cent
 FT identity with FKBP and comprises 4 zones of
 FT insertions or deletions"
 FT 254..267
 FT /label= hinge
 FT 268..372
 FT /label= HBI-III
 FT /note= "globular domain"
 FT 373..458
 FT /label= HBI-III
 FT /note= "C-terminal
 FT /label= CAM"
 FT 33..39
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 53..61
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 66..68
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"

FT region
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 88..97
 FT /note= "alpha-helix"
 FT 103..106
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 128..137
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 150..156
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 170..178
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 182..184
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"
 FT 188..192
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 202..211
 FT /note= "alpha-helix"
 FT 217..220
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 243..252
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 269..275
 FT /label= A1
 FT /note= "antiparallel beta-pleated sheet"
 FT 289..297
 FT /label= A4
 FT /note= "antiparallel beta-pleated sheet"
 FT 301..303
 FT /label= A5
 FT /note= "antiparallel beta-pleated sheet"
 FT 307..311
 FT /label= beta
 FT /note= "supplementary beta-pleated sheet"
 FT 321..330
 FT /note= "alpha-helix"
 FT 335..338
 FT /label= A2
 FT /note= "antiparallel beta-pleated sheet"
 FT 362..371
 FT /label= A3
 FT /note= "antiparallel beta-pleated sheet"
 FT 399..414
 FT /note= "site of interaction with calmodulin"
 FT 441..458
 FT /note= "used to generate polyclonal antibodies"
 FT region
 FT binding_site
 FT 16-SEP-1993.
 FT 04-MAR-1993; F00219.
 FT (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 FT Baulieu E, Callebaut I, Chambraud B, Ledreau M, Massol N;
 FT WPI: 93-303460/38.
 FT N-PSDB: Q48515.
 FT New protein forming complex with heat shock protein - also
 FT binding immuno-suppressors, etc., and corresp. nucleic acid,
 FT antibodies etc., useful e.g. for detecting tumours, treating
 FT auto-immune disease, etc.
 PS Claim 1; Fig 1; 43pp; French.
 CC The HBI protein is able to complex chaperone protein hsp90, even
 CC when the chaperone protein is part of a hetero-oligomer with other
 CC proteins. Hsp90 can bind to steroid hormone receptors, vitamin D,
 CC and Tyrosine Kinases of viral oncogenes; HBI is thus useful for the

Fri Sep 3 14:48:31 1999

US-09-120-044-3-COPY.rag

Page 8

CC study, prevention or treatment of diseases associated with
CC dysfunction of proteins which form complexes with hsp90, e.g.,
CC ricketts, cancer, dioxin poisoning or autoimmune disease. HBI can
CC also be used in immunosuppressor research.
SQ Sequence 458 AA;

dysfunction of proteins which form complexes with hsp90, e.g.

CC also be used in immunosuppressor research.

Sequence 458 AA;

Sequence 458 AA;

Query Match	3.08;	Score 96;	DB 8;	Length 458;
Post-read Similarity	33.38;	Prod No	6.61e+01;	

Best Local Similarity 22.28; Pred. No. 6.61e+01;

Matches	20;	Conservative	19;	Mismatches	49;	Indels	2;	Gaps	2;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

Db 221 lkpysafnagkekfqippyaelkyevh1ksfekakesw emsseeklegsaivkergtv 279

176 LDIDNSVHSGEKQIQIVNKKQIYYTIVSDAVKNPQGVFQDVTVTEDLKQRGISAEKPLV 235

Db 280 yfkegky-kqalqykktivswleyessfs 308

QY 236 YISXVAYGRQVYLKLETTSSXSXEVEAAFEA 265

Search completed: Fri Sep 3 13:13:36 1999
Job time : 110 secs.

Job time : 110 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Run on: Fri Sep 3 13:10:39 1999; Maspar time 20.83 Seconds
Tabular output not generated. 905.977 Million cell updates/sec

Title: >US-09-120-044-3-COPY
Description: (1-471) from dev1120044.pep
Perfect Score: 3234
Sequence: 1 MANKAVNDFILAMNYDKKLT.....TISWGTTLYPQVEDKEND 471

Scoring table: PAM 150
Gap 11

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir60
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 50.326; Variance 110.458; scale 0.456

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3232	99.9	471	2	A28568	pneumolysin - Strepto	0.00e+00
2	1600	49.5	500	2	B43577	perfringolysin O prec	2.15e-270
3	1543	47.7	529	2	A43505	listeriolysin O precu	1.85e-259
4	1536	47.5	529	2	S24231	listeriolysin O precu	4.06e-258
5	1526	47.2	538	2	S22341	ivanolysin precursor	3.35e-256
6	1477	45.7	530	2	S22340	seeligeriolysin - lls	8.14e-247
7	1461	45.2	501	2	A37858	alveolysin - Bacillus	9.40e-244
8	1408	43.5	571	2	A43507	streptolysin O precu	1.29e-233
9	1329	41.1	485	2	I39863	hemolysin - Bacillus	1.55e-218
10	464	14.3	112	2	S47298	sulfolysin - Streptococ	2.28e-57
11	214	6.6	96	2	S47297	sulfolysin - Streptococ	3.28e-15
12	117	3.6	527	2	S03974	amine oxidase (flavin	1.18e-01
13	112	3.5	748	2	S59327	hypothetical protein	4.55e-01
14	107	3.3	218	2	F64461	hypothetical protein	1.68e+00
15	106	3.3	285	2	S60947	hypothetical protein	2.17e+00
16	107	3.3	462	1	S00552	mitochondrial protein	1.68e+00
17	107	3.3	501	2	A69415	conserved hypothetical	1.68e+00
18	108	3.3	511	2	S58322	nucleolar protein NOP	1.30e+00
19	107	3.3	1228	2	I40468	surface layer protein	1.68e+00
20	106	3.3	1361	3	T03415	S-layer protein - Cam	2.17e+00
21	107	3.3	5762	2	A41819	proline-rich peptides	1.66e+00
22	104	3.2	187	2	C71140	hypothetical protein	3.61e+00
23	103	3.2	303	2	S70118	probable membrane pro	4.64e+00

24	102	3.2	365	2	F71665	hypothetical protein	5.96e+00
25	105	3.2	426	2	A70444	flagellar hook associ	2.81e+00
26	102	3.2	446	2	E69899	L- amino acid oxidase	5.96e+00
27	104	3.2	493	2	S07375	flagellin H-1 - Salmo	3.61e+00
28	105	3.2	505	2	S09638	flagellin - Salmone	2.81e+00
29	105	3.2	514	1	DWBRTT	threonine dehydratase	2.81e+00
30	102	3.2	756	2	C64236	protein V (TcrV) homo	5.96e+00
31	102	3.2	1035	2	A64686	cell division inhibi	5.96e+00
32	101	3.1	259	2	B69113	hypothetical protein	7.63e+00
33	101	3.1	290	2	S56569	probable aminotransf	7.63e+00
34	100	3.1	362	2	B70876	hypothetical protein	7.63e+00
35	101	3.1	418	2	G71952	glycine hydroxymethyl	9.76e+00
36	100	3.1	427	2	S75210	arginine decarboxylas	9.76e+00
37	100	3.1	502	2	U02341	hypothetical protein	9.76e+00
38	100	3.1	513	2	A65249	hypothetical protein	9.76e+00
39	100	3.1	666	2	T03090	hypothetical protein	9.76e+00
40	101	3.1	722	2	S64492	hypothetical protein	9.76e+00
41	100	3.1	878	2	F64425	valine--tRNA ligase (9.76e+00
42	101	3.1	1035	2	D71831	cation efflux system	7.63e+00
43	101	3.1	1123	1	WMBEH7	U537 protein - human	7.63e+00
44	99	3.1	1681	2	S42369	ClaHtrn heavy chain	1.24e+01
45	100	3.1	1744	2	JH0720	tanabin - African cla	9.76e+00

ALIGNMENTS

RESULT 1
ENTRY A28568 #type complete
TITLE pneumolysin - Streptococcus pneumoniae
ORGANISM #formal name Streptococcus pneumoniae
DATE 19-Nov-1988 #sequence-revision 19-Nov-1988 #text-change 21-Aug-1998

ACCESSIONS A28568; S12829
REFERENCE A28568
#authors Walker, J.A.; Allen, R.L.; Falmagne, P.; Johnson, M.K.;
Boulois, G.J.

#journal Infect. Immun. (1987) 55:1184-1189
#title Molecular cloning, characterization, and complete nucleotide
sequence of the gene for pneumolysin, the
sulphydryl-activated toxin of Streptococcus pneumoniae.

#cross-references M01D:87193109
#accession A28568
#molecule-type DNA

REFERENCE S12829
#residues 1-471 #label MAL

#authors Mitchell, T.J.; Mendez, F.; Paton, J.C.; Andrew, P.W.;
Boulois, G.J.

#journal Nucleic Acids Res. (1990) 18:4010
#title Comparison of pneumolysin genes and proteins from
Streptococcus pneumoniae types 1 and 2.
#cross-references M01D:90326546
#accession S12829
#molecule-type DNA

#status Preliminary
#molecule-type DNA

CLASSIFICATION #presidues 1-471 #label MIT
FEATURE ##cross-references EMBL:X52474; M01D:947403; PID:947404
#superfamily dipeptide transport protein

SUMMARY 2-471
#product pneumolysin #status predicted #label MAT
#length 471 #molecular-weight 52899 #checksum 8274

Query Match 99.9%; Score 3232; DB 2; Length 471;
Best Local Similarity 97.9%; Pred. No. 0.00e+00;

Matches 461; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Db	1	MANKAVNDFILAMNYDKKLTTHQGESIENRFKEGNOUPDEEVVERKKRSISTSTSI	60
Oy	1	MANKAVNDFILAMNYDKKLTTHQGESIENRFKEGNOUPDEEVVERKKRSISTSTSI	60
Db	61	SVTATNDSRLYPGALLVDETLLENNPTLLAVRADMTYSIDLPGLASSDSFLOVEDPSN	120
Oy	61	XVTATNDSRLYPGALLVDETLLENNPTLLAVRADMTYSIDLPGLASSDSFLOVEDPSN	120

Db 121 SSVGAANDLAKWHQDYGQVNNVPAKMOYEKITAHSMDLKVKGSGDEPKTGNLSIDIF 180
121 SSVGAANDLAKWHQDYGQVNNVPAKMOYEKITAHSMDLKVKGSGDEPKTGNLSIDIF 180
QY 121 SSVGAANDLAKWHQDYGQVNNVPAKMOYEKITAHSMDLKVKGSGDEPKTGNLSIDIF 180
Db 181 NSVHSGEKOIYVFKQIYITVSDAVKPNQDVFQDVTYVEDLKQKISAEPLVYISXV 240
181 NSVHSGEKOIYVFKQIYITVSDAVKPNQDVFQDVTYVEDLKQKISAEPLVYISXV 240
QY 181 NSVHSGEKOIYVFKQIYITVSDAVKPNQDVFQDVTYVEDLKQKISAEPLVYISXV 240
Db 241 AYGGQVYLKLETTESKDSVEAEAFALIKGVAPQTEMKOILNTEKAVIILGDPSSGA 300
241 AYGGQVYLKLETTESKDSVEAEAFALIKGVAPQTEMKOILNTEKAVIILGDPSSGA 300
QY 241 AYGGQVYLKLETTESKDSVEAEAFALIKGVAPQTEMKOILNTEKAVIILGDPSSGA 300
Db 301 RVATGVDMVEDLIQEGSRFTADHGPGLISYTTSTFLDNVVAFFQNSTDYETKVTAYRN 360
301 RVATGVDMVEDLIQEGSRFTADHGPGLISYTTSTFLDNVVAFFQNSTDYETKVTAYRN 360
QY 301 RVATGVDMVEDLIQEGSRFTADHGPGLISYTTSTFLDNVVAFFQNSTDYETKVTAYRN 360
Db 361 GDLLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 420
361 GDLLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 420
QY 361 GDLLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 420
Db 421 LSVKIRECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIYPOVEDKVEN 471
421 LSVKIRECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIYPOVEDKVEN 471
QY 421 LSVKIRECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIYPOVEDKVEN 471
RESULT 2 B43577 #type complete
ENTRY perfringolysin O precursor - Clostridium perfringens
TITLE #formal_name Clostridium perfringens
ORGANISM 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change
DATE 13-Sep-1998
ACCESSIONS B43577; A34951; A60922
REFERENCE A43577
#authors Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
#journal Infect. Immun. (1991) 59:137-142
#title An upstream regulatory sequence stimulates expression of the
perfringolysin O gene of Clostridium perfringens.
#cross-references MUID:91099951
#accession B43577
#molecule_type DNA
#residues 1-500 #label SHI
#cross-references GB:M81080; NID:9144885; PID:9144886
#note Translation of the nucleotide sequence is not complete
REFERENCE A34951
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3235-3240
#title Nucleotide sequence of the gene for perfringolysin O
(theta-toxin) from Clostridium perfringens: significant
homology with the genes for streptolysin O and pneumolysin.
#cross-references MUID:89032623
#accession A34951
#molecule_type DNA
#residues 1-125, 'EA', 129-500 #label TWE
#cross-references GB:M36704; NID:9144883; PID:9144884
REFERENCE A60922
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3228-3234
#title Cloning and expression in Escherichia coli of the
perfringolysin O (theta-toxin) gene from Clostridium
perfringens and characterization of the gene product.
#cross-references MUID:89032622
#accession A60922
#molecule_type protein
#residues 29-45 #label TW2
#experimental_source ATCC 13124
GENETICS pfo; pfoA
#gene
FEATURE 1-28 #domain signal sequence #status predicted #label SIGV
29-500 #product perfringolysin O #status experimental #label
MAR
SUMMARY #length 500 #molecular-weight 55799 #checksum 4821

Query Match 49.5%; Score 1600; Db 2; Length 500;
Best Local Similarity 46.5%; Pred. No. 2,15e-270;
Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;
Db 35 NOSIDSGISSLSTYRNNEVLAANGKISFVPEKERRKGNKFIYVERKRLTSPVDISI 94
35 NOSIDSGISSLSTYRNNEVLAANGKISFVPEKERRKGNKFIYVERKRLTSPVDISI 94
QY 3 NKAANDFLIANNKKKLLTHQGESIENRFKESGNQJDEFEVXERRKRLSTINTSDIXV 62
3 NKAANDFLIANNKKKLLTHQGESIENRFKESGNQJDEFEVXERRKRLSTINTSDIXV 62
Db 95 IDSVNDRTYPPGALOLADKAEFENRPTIIMYKRKPININIDPLGKJGNS-1KYDDPTYGK 153
95 IDSVNDRTYPPGALOLADKAEFENRPTIIMYKRKPININIDPLGKJGNS-1KYDDPTYGK 153
QY 63 TATNDSRLYPPGALLVDEPTLKENNPITLADRAMYTSIDLPGIASSDSPLOVEDPENSS 122
63 TATNDSRLYPPGALLVDEPTLKENNPITLADRAMYTSIDLPGIASSDSPLOVEDPENSS 122
Db 123 VSGAIDELVSKWNEKYSSTHTLPARTQYSESVMYSKQISALANNAKVLNLSIGDENA 213
123 VSGAIDELVSKWNEKYSSTHTLPARTQYSESVMYSKQISALANNAKVLNLSIGDENA 213
QY 123 VSGAIDELVSKWNEKYSSTHTLPARTQYSESVMYSKQISALANNAKVLNLSIGDENA 213
Db 214 VANNKKMILAIYKQIETFTSADLPKPNSDLPDSDVFNFLKQGVNENAPPLMVSVAV 273
214 VANNKKMILAIYKQIETFTSADLPKPNSDLPDSDVFNFLKQGVNENAPPLMVSVAV 273
QY 183 VHSGEKOIYVFKQIYITVSDAVKPNQDVFQDVTYVEDLKQKISAEPLVYISXV 242
183 VHSGEKOIYVFKQIYITVSDAVKPNQDVFQDVTYVEDLKQKISAEPLVYISXV 242
Db 274 GTTIVKLETTSSKDVQAAFKALIKNDIKNSQOYKDIYENSFTAVVLDGDAQHNV 333
274 GTTIVKLETTSSKDVQAAFKALIKNDIKNSQOYKDIYENSFTAVVLDGDAQHNV 333
QY 243 GRQVYLKLETTESKDSVEAEAFALIKGVAPQTEMKOILNTEKAVIILGDPSSGARV 302
243 GRQVYLKLETTESKDSVEAEAFALIKGVAPQTEMKOILNTEKAVIILGDPSSGARV 302
QY 334 YTKDFEIRKVIKDNATESTRNPAVPISTYSVFLKDNVAAVHNKTDYIETTSTESK 393
334 YTKDFEIRKVIKDNATESTRNPAVPISTYSVFLKDNVAAVHNKTDYIETTSTESK 393
QY 303 VVGKVDVEDLIQEGSRFTADHGPGLISYTTSTFLDNVVAFFQNSTDYETKVTAYRN 362
303 VVGKVDVEDLIQEGSRFTADHGPGLISYTTSTFLDNVVAFFQNSTDYETKVTAYRN 362
Db 394 ILLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 422
394 ILLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 422
QY 363 ILLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 422
363 ILLDHSAGVAAQYIITWELSYDHQKEVLPKAMDNGODLTAFTTISIPKGNVRN 422
Db 454 IKARECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIY 493
454 IKARECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIY 493
QY 423 VKIRECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIY 462
423 VKIRECTGLAMEMWRTVYEKTDPLVRRKRTISMGTTIY 462
RESULT 3 A43505 #type complete
ENTRY listeriolysin O precursor - Listeria monocytogenes
TITLE #formal_name Listeria monocytogenes
ORGANISM 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change
DATE 17-Mar-1999
ACCESSIONS A43505; S05306; A47606; S12400; A61079
REFERENCE A43505
#authors Mengaud, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.;
Geoffroy, C.; Gicquel-Sanzey, B.; Baquero, F.; Perez-Diaz,
J.C.; Coesart, P.
#journal Infect. Immun. (1988) 56:766-772
#title Expression in Escherichia coli and sequence analysis of the
listeriolysin O determinant of Listeria monocytogenes.
#cross-references MUID:88153053
#accession A43505
#status preliminary
#molecule_type DNA
#residues 1-529 #label MEN
#cross-references GB:M24199; NID:9149652; PID:9149653
#note this sequence is derived from a strongly hemolytic
strain, serotype 1/2c
REFERENCE S05306
#authors Domann, E.; Chakraborty, T.
#journal Nucleic Acids Res. (1989) 17:6406
#title Nucleotide sequence of the listeriolysin gene from a Listeria
monocytogenes serotype 1/2a strain.
#cross-references MUID:89366684
#accession S05306
#molecule_type DNA
#residues 1-529 #label DOM
#cross-references EMBL:X15127; NID:944106; PID:944107
#experimental_source strain Egd
#note this sequence is derived from a weakly hemolytic strain,
serotype 1/2a
REFERENCE A47606

Query Match	Best Local Similarity	Matches	Score	DB 2	Length	529
61	1DKY1QGLDYNKNNLVYHGDAVTNVPPRKGKDGNGEYIVVEKKKKSINQNNADIQYVNA	200	47.7%	117	148	1
6	VNDFITLAMYDKKKLLTHQGESIENRFKQEGNOJLPDEFVVAERKKRSLSTWTSIDIXVAT	200	42.9%	117	148	1
121	ISSLTYPGALYKANSELYENQPDVLPVKRDSLTLSIDLPKMTNODKNIVYKNAIKSNVNN	200	47.7%	117	148	1
66	NDSRLYPGLLVDDTLTKENNPSTLLAVDRAAPPTYSIDLPGLASSDSFLQVDPSPNSVYG	200	47.7%	117	148	1
181	AVNTLYVERANEKYAAQAYPRVSAKIDYDDEMAVSESQLAKFGTAFKAVNNSLVNFGAIS	200	47.7%	117	148	1
126	AVNDLLAKKHQYQGVN-NVPAKRYEKRTIASHMQLKVKFGSDPEKGTGNSLIDFNSTVH	200	47.7%	117	148	1
241	EKKQGEVVISFKQIYYNVNNEPTRPSPREFGKAVTKEQLQALGVNAENPPAYISSVAYGR	200	47.7%	117	148	1
185	SEKQIOIYNKQIYYTVSDAVKKNPGGVFOPTVEDLKORGISAERPLVYISXVAYGR	200	47.7%	117	148	1
301	QVYLKLTSTSHSTKKAADFVAIVSKSVSGDELNIINKNSFRKAVIYGGSKAKDEVOIJD	200	47.7%	117	148	1
245	QVYLKLETTSSXSEVEAAEALIKGVKAPAPQTEMKQILDNTEVKAIVLIGGPPSSGARVY	200	47.7%	117	148	1
361	GNLGLDRDLIRLKGAFNFRPEVPIATYTNFNLKDELAVIKNSSEYIEFTSKAAYDGKIN	200	47.7%	117	148	1
305	KGAVNEDDIDGSGSRRTADHPRLPISYTSFLRDVNVATFQKSTDYVELTKYATAYNGDL	200	47.7%	117	148	1
421	IGHSGGYAOFNISWDEVNYPDEGNEIVQHKWMSSENNKSKLAHFTSSIYLPQGNANINY	200	47.7%	117	148	1
365	LHSGGAYVAYQYIYWNELSYDHQGEVLLPCKAMDRNGODLLAHFTSSIPLKQGNVNLVSK	200	47.7%	117	148	1
481	AKRGTGLAEMWRVTYIDDNRLPLVKNRNTSINGTLLPLPYSSKXVNN	200	47.7%	117	148	1
425	IRRCTGLAEMWRVTYIEKTDPLVKKRRTISIGTLLYPQVEDKVEN	200	47.7%	117	148	1

ORGANISM	12067)	#formal_name	Listeria monocytogenes
DATE	22-Nov-1993	#strain	12067
DATE	13-Sep-1998	#sequence_revision	10-Nov-1995
DATE		#text_change	
ACCESSIONS	S24231		
REFERENCE	S24230		
#authors	Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.		
#journal	Infect. Immun. (1991) 59:3945-3951		
#title	Listeria monocytogenes isolates can be classified into two major types according to the sequence of the listeriolysin gene.		
#accession	MUID:92040062		
#status	524231		
#molecule_type	DNA		
#residues	1-529		
#cross-references	EMBL:X60035; NID:g44110; PID:g44112		
#experimental_source	strain 12067, serotype 4b		
#note	the nucleotide sequence was submitted to the EMBL Data Library, June 1991		
GENETICS			
#gene	lisa		
CLASSIFICATION	lisa		
FEATURE	virulence factor		
1-75			
26-529			
SUMMARY			
Query Match	47.5%;	Score 1536;	DB 2; Length 529;
Best Local Similarity	42.7%;	Pred. No. 4,066-258;	
Matches	199;	Conservative 117;	Mismatches 149; Indels 1; Gaps 1
Db	61	IDKYTGDLNNNNLVHGDATVWPVRKQYKQGNEXIVEEKKKKSINONNADIQVNA	120
Qy	6	VNDFLIANNYDKRKLITLQGESEIERKFKBEGNQLPDEVVAXERKRRLSTISDIXYAT	65
Db	121	ISSLTYPGALVKSANSELYENOPDVLVPRKDSLTLSIDLPGMTNODNKIVYNAKTSNVNN	180
Qy	66	NOSRLYPGALVDEFLXENNPNTLLAVDRAPMTYSIDLPGIASSDSFLQVEDPSSSVRG	125
Db	181	AVNTLVERNNEYAQAIPYNSAKIDYDDEMAVSESQILAAGTAFAKAVNSLNANFGAIS	240
Qy	126	AVNDLAKKHOHOGYGVN-NVPAKQYKEKITAHSMDQLKVGKSGFEKTSGLSDIDENSVAH	184
Db	241	EGKMOEYEFSPQIYYNVNNEPTSPSRFFGKAVKEQOLAQVNAENPPIAYISVAVGR	300
Qy	185	SGEKQIQIYNKQIYYTYSVDVAKNPGDFQDPTVYEDLQKRGISAERPPLYISYAVAGR	244
Db	301	OYVTLKSTNSHTKYKAAFDAVASGKSVSGDELNTIINKSSFKAVIYVGSAGKDEVOIID	360
Qy	245	OYVTLKSTNSHTKYKAAFDAVASGKSVSGDELNTIINKSSFKAVIYVGSAGKDEVOIID	360
Db	361	GNLGLRLDLKKGAFNRPETPVPAITTNPKDELAVIKNSSEYIETTSKATYDGIN	420
Qy	305	GKDWEDIEDIOGSRFADHPPLPSYTSFSLRDVVVAITFONSDIYEITKYTAARNQDL	364
Db	421	IDHSGYVQNPINSMDEIYDEGNEIYOHKWSNNKSKLAHFSSYLLPGNARNINWY	480
Qy	365	LDHSGYVQNPINSMDEIYDEGNEIYOHKWSNNKSKLAHFSSYLLPGNARNINWY	480
Db	481	AKECTGLAEMWRTYIDRNLPLVKNRMSISWGLTLYPKYSNVN	526
Qy	425	IRECTGLAEMWRTYIEKTDLPVLRKRTISWGLTLYPQVEDKVEN	470
RESULT	5		
ENTRY	S22341	#type complete	
TITLE	ivanoliysin precursor - Listeria ivanovii		
ORGANISM	07-Apr-1994	#sequence_revision	07-Apr-1994
DATE	21-Aug-1998	#text_change	

ACCESIONS	S22341: S36683
REFERENCE	S22340
#authors	Haas, A.; Dumbeky, M.; Kreft, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	listeriolysin genes: complete sequence of <i>ilo</i> from <i>Listeria</i> Ivanovii and of <i>iso</i> from <i>Listeria seeligeri</i> .
#cross-references	MUID:92182018
#accession	S22341
##molecule_type	DNA
##residues	1-528 ##label HAS
##cross-references	EMBL:X60461
##note	the authors translated the codon ACA for residue 331 as Val
REFERENCE	S36683
#authors	Kreft, J.
#submission	submitted to the EMBL Data Library, July 1991
#accession	S36683
##molecule_type	DNA
##residues	1-319, 'T', 321-528 ##label KRE
##cross-references	EMBL:X60461
GENETICS	
#gene	<i>ilo</i>
CLASSIFICATION	#superfamily dipeptide transport protein
FEATURE	
1-24	
25-528	#domain signal sequence #status predicted #label SIG\
SUMMARY	#product <i>listeriolysin</i> #status predicted #label MAT
	#length 528 #molecular-weight 58511 #checksum 6674
Query Match	47.2%; Score 1526; DB 2; Length 528;
Best Local Similarity	42.9%; Pred. No. 3.35e-256;
Matches	200; Conservativity 118; Mismatches 147; Indels 1; Gaps 1;
Db	60 IDQYIOGLDYDKRNLLVYDGEAVNVNPKAKYKGGNOIVYERKKKSINONNADIOVNS 119
QY	6 VNDFTLANVYDKKLLTHQGESIENRFKBEQNPDERVYERKKRSLSTSTSDIXTAT 65
Db	120 LASLTYPALVKANSELEVNOPDYLPVKRDSYTLSTIDPGAVNHNDNFIYQNAKSNIND 179
QY	66 NDSRLTPGALLVVDETLKENNPFTLLADRAEMTYSIDLPGIASSDSFLQVEDPSSVVG 125
Db	180 GVNTLVDPNNKRYSEEEYNISAKIDYDQEMAYSSQIARVKAAGFAKAVNNSLWNGEAS 239
QY	126 AVNDLLAKWHDDYQGVN NVPARQYKEITHTSHSQLKVFSGSFETGSLSDIDENSVH 184
Db	240 EGYQVEEYINFKQIYYVNVNEPTSPSRFEKSVTKENLOALGVNAENPPAYTSSVAYGR 299
QY	185 SGEKOIQIVNNKQIYYVNVSDAVNPGVDFODTVEDLKGORGISABRPVYTSSVAYGR 244
Db	300 DIFKLTSSSTSVKKAFAADAFKSGYKGGTLENTIIONASFAYVYGSADAVEIID 359
QY	245 QYIATLETTSSKEVEAFALIGVAVAPOTERKQILDNTIEKAAVILGDPSSGARVYT 304
Db	360 GDSLKLDIILKOGANFEKDKNPVBIATYTNFLKNOIAVAVKNNSEYIETTSKAYSDGKIN 419
QY	305 GKVMVEDLLDEGSRFADHBPGLISYTTSLPBDNVAVATFQNSDYETKVTAYRNGDLL 364
Db	420 LDHSGAVARBNVWDVSDYDANSEVENYERKKWSBENKDKLAHTTSTIYIPGAKRNINH 479
QY	365 LDHSGAVAYAAQIYYTNWELSDIHQSEKVELTPAPAMRNQODTLAHTTSTIPLKGVNRLSVK 424
Db	480 AKECTGLAMWVRTVDDRNLLPLVKNKNVCIWGTTLTPAYSDVDN 525
QY	425 IRECTGLAMWVRTVYKTDLPVKKRTISIWGITTLTPQVEDKVEN 470
RESULT	6
ENTRY	S22340
TITLE	seeligeriolysin - <i>Listeria seeligeri</i>
ORGANISM	#formal_name <i>Listeria seeligeri</i>
DATE	22-Nov-1993 #sequence_revision 01-Dec-1995 #text-change 21-Aug-1998
ACCESIONS	S22340
REFERENCE	S22340

#authors	Haas, A.; Dumbsky, M.; Kreff, J.
#journal	Biochim. Biophys. Acta (1997) 1130:81-84
#title	Listeriolysin genes: complete sequence of llo from <i>Listeria</i> Ivanovii and of lso from <i>Listeria seeligeri</i> .
#cross-references	WUID:92182018
#accession	S22340
#status	Preliminary
#molecule_type	DNA
#residues	1-530 ##label HMA
#cross-references	EMBL:X60462; NID:g44144; PID:g44145
#note	The authors translated the codon GCC for residue 287 as Pro
CLASSIFICATION	#superfamily dipeptide transport protein
SUMMARY	#length 530 #molecular-weight 59181 #checksum 4523
Query Match	45.7%; Score 1477; DB 2; Length 530;
Best Local Similarity	42.5%; Pred. No. 8,14e-247;
Matches 198;	Conservative 111; Mismatches 156; Indels 1; Gaps 1
Db	62 INKYLNDKNSILVYQGEAVTVPPKGYKDGSEIVVEKKKKGINONNADISVINA 121
Qy	6 VNDEILAMNDKKLLKHGGESLENRFKKEGNOQDPDEVVYERKKRSLNSTDIVATF 65
Db	122 ISSLTYGALYKANRELVENQPNVLPVKRDLITLSVDLPGMTKDKNKFYKAPFTKSNVN 181
Qy	66 NDSRLYGGALLVDEFTLKENNPPTLLAVRAPRMTSIDPLGSSDSDFLQVEDPSNMSVVG 125
Db	182 AVNTLVERMADKSKAPYINPKIDYSEDMAYSESOLIAKGTAFKAVNNSLVNVEATS 241
Qy	126 AVNDLLAKWHDQGVQNV-NVPRARQYEKLTASHMQLKAFKFGSPSEFTGNSLSDIDNSVH 184
Db	242 DGAVOEEFVKOIIYVINVNEPTSPSKFSGSVTKQDLALGVNAENPPATISSVAYGR 301
Qy	185 SGEKQIDIVAKKQIYYTVSGDVAKKPGDVFQDVTVEDLKQRGISARPLVYISVAYGR 244
Db	302 QVYKLTSSSSHKVKTAFKAFAMSKSGKGVDELTIINXKSPFRAYVIGSAAEEVEIID 361
Qy	245 QVYIKLETTISXSEVAAFAELALIKGVKAPQTEKQILDITVEKAAVILGGDPSGARGVT 304
Db	362 GNIGELDLIKSGSYDRENQVPISITYTNFLKNDLAVVAKNNSEVIEETTSKSYTGKIN 421
Qy	305 GKVDWVEDLIQESRFTADHPPLPISYTSFLRDNVAVATFQNSDYEKRVIAIYARNGDLL 364
Db	422 IDHSGGVAFNLSMDEVSVDNGEIEIVHKKGEKNTKSLAFTSSITLPGNARNINITY 481
Qy	365 LDHSGAVVAQYITWELSYDQGEKVELTPFAMRNODTLAHTTSTIPLKGAVNRLSVK 424
Db	482 ARECTGLFEMWRTVIDDRNLPLVKNRVISMGTLTPRHSNVND 527
Qy	425 IREOTGLAMEMWRTVEKTDPLVAKRRTISMGTLTPQVEDKVEN 470
RESULT	7
ENTRY	A37858 #type complete
ORGANISM	alveolysin - <i>Bacillus alvei</i>
DATE	28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 09-Sep-1997
ACCESSIONS	A37858
REFERENCES	A37858
#authors	Geoffroy, C.; Mengaud, J.; Alouf, J.E.; Cossart, P.
#journal	J. Bacteriol. (1990) 172:7301-7305
#title	Alveolysin, the thiol-activated toxin of <i>Bacillus alvei</i> , is homologous to listeriolysin O, pertussin toxin O, pneumolysin, and streptolysin O and contains a single cysteine.
#cross-references	MUID:91072294
#accession	A37858
#status	Preliminary
#molecule_type	DNA
#residues	1-501 ##label GEO
#cross-references	GB:M62703; NID:g142472; PID:g142473
#summary	#length 501 #molecular-weight 55268 #checksum 8775

Query Match	45.2%: Score 1461; DB 2; Length 501;
Best Local Similarity 41.7%: Pred. No. 9, 40e-244;	
Matches 189; Conservative 121; Mismatches 142; Indels 1; Gaps 1;	
Db	44
IAGLANTNRNNEVLAIOGDISSEFVPEKGIOSNGKEFIVERDKSLTSPYDISIVDSITNR	103
10 ILAMVQDKRLTLHOGESIENRFKCEQNLDPDEFVXERKRKSLSTNTSDIXVATANDSR	69
Db	104
TYRPAIOLAKKDPADNOPSLVMAARKPLDISIDLPGLKNEHT-ISTQNVNGVSSAIDQ	162
70 LYPALLLVDETLEKNPPTLLAADRAPMTYSIDLPGIASSDSFLQYEDPSNSSVRAVND	129
Db	163
LVSTWGEKSYSTHTPLRLQYAESMYSOMQISSALLVNAKVLNGLTIDFNANVANGCK	222
Qy	130
LIAMHHDQYQGVNNVPARKXQYEKTIASHMQLKVKGSDPEKGNLDDIFNSVHSGEQ	189
Db	223
VMAAVYKQIFVTVSAGLPPNPSDLFDDSVTFPAELARKGVSENPAPLWVNAVGRITVYK	282
Qy	190
IQIYNXKQIYTVTVSDAVAKNPGDVFQDTVEDLKGQISAEERPLVYISXVAGROYVTK	249
Db	263
LETTSKENDVQTAARKLLNNPSTQASQYKDIENSFTAVVYGGAQHINQVYTDENV	342
Qy	250
LETTSKXKEVEAAEALIKGVKAPQTEWKQILLDNEVKNVILGGPOSSARVATKVM	309
Db	343
IOSVYIKNAQFSKNRPAPISYTSVFLKDNSIAAVHNTEYIEKTTEYSKGIKIDHSG	402
Qy	310
VEDLIOGSRRTAHPGLPISYTSFLRDVNAVATFQNSTDYVELTKYATANGDGLLDHSG	369
Db	403
AYVAQFEVYWDSESYDADGQETVTRKSMQDGNMRDSAHFSTEIPLPNAKINIRIFARECT	462
Qy	370
AYVAQYITVNETLISYDHQGEVTLPRKAMRGODTLAHTFSTISPLKGNVAVLSVKIRECT	429
Db	463
GLAEMWMTVYDEVNVPASIDANSINGTLYP	495
Qy	430
GLAEMWMTVYEKTDLPVKKRTISWGTTLYP	462
RESULT	8
ENTRY	A43507
TITLE	streptolysin O precursor - Streptococcus pyogenes
ORGANISM	Formal_name Streptococcus pyogenes
DATE	28-Oct-1992 #sequence revision 28-Oct-1992 #text change 09-Sep-1997
ACCESSIONS	A43507
REFERENCES	A43507
#authors	Kehe, M.A.; Miller, L.; Walker, J.A.; Boulinois, G.J.
#journal	Infect. Immun. (1987) 55:3228-3232
#title	Nucleotide sequence of the streptolysin O (SLO) gene: structural homologs between SLO and other membrane-damaging, thiol-activated toxins.
#cross-references	MUID:88057628
#accession	A43507
#status	preliminary
##molecule_type	DNA
##residues	1-571 #label KEH
##cross-references	GB:M18638; NID:G153810; PID:G153811
SUMMARY	#length 571 #molecular_weight 63638 #checksum 73
Query Match	43.5%: Score 1408; DB 2; Length 571;
Best Local Similarity 41.8%: Pred. No. 1, 28e-233;	
Matches 191; Conservative 104; Mismatches 161; Indels 1; Gaps	1;
Db	109
INDKYSILNNELEVAENGEIEFVPEKGVKADKRVTERKKNNINTPVDISITDS	168
Qy	6
VNDPLIANNYDKKLLTHOGESIENRFKKEGNQLPDEFVYERKKRSLSTNTSDIXVAT	65
Db	169
VTRDTYPAALOLANKGTEENKPDVAVYTRNPKOIHIDLPNGDKRAT-VEVNDPTVANST	227
Qy	66
NDSRLYVPGALLVDEFLXENNPTLLAVRAPMTYSIDLPGIASSDSFLQYEDPDSNVYRG	125
Db	228
AIDNLTNOMHNDYSGNLTLPATQYTESMYSKSLDIALVANSKILDTGLIDFKSISK	287

Qy	126	AVNDLLAKKHNDYGGVNNVPARXQYERKTLTASHMEQJLKVKGSGFEKTSGLIDENSVHS	185
Db	288	GEKKVMAIYAKQIFTTVSANLPNNPADYEDKSTYFKLQKGSYNAAPLPFVSNAVYERT	347
Qy	186	GGGQIQIVAKKQIYYTVSYDAKMGKGVDFQDTYVEDLKRGISAEPRLPYISXVAAYGRQ	245
Db	348	VFKLETFSSKSNDEVAEFAALAKGVDYNTNGKYSDILENSSFVAVLGDAAEHKKVYTK	407
Qy	246	YVTKLETFSSKSEVEEAPFALIKGVKVAPOYEMKQILDNTYKVAVILGGDPSSGARVYTG	305
Db	408	DFVDYRNVIKDNATESRKNPAYISYTSVFLKNNKLAGVNRNTEYVETTSYETSQINL	467
Qy	306	KVMDVEDLIEGSRFPADHPGLDIPYTSFSLRDVNAVTFQNSDYETVETVYARNGDILL	365
Db	468	SHGAAVYAQVAILMDINVDKGRKVIRKRDNNWAKTSPESTYVILPGASNRNIRIA	527
Qy	366	DHSGAVYAQIYYIWNELSTHDHGKVEVLTPKAMDNGODLTAHFITSLIPKGNRNLVSKI	425
Db	528	RECTGLAMEWRKVIDERDVKLSKINNVINISGSTLSP	564
Qy	426	RECTGLAMEWRKVYEXKIDLPYKRRTISIMGTTLYP	462
RESULT	9		
ENTRY		I39863	#type fragment
TITLE		hemolysin - Bacillus cereus (fragment)	
ORGANISM		#formal_name Bacillus cereus	
DATE		19-Jul-1996	#sequence_revision 19-Jul-1996
ACCESSIONS		I39863	
REFERENCE		I39863	
#authors		Yutsudo, T.; Okumura, K.; Iwasaki, M.; Hara, A.; Kakitani, S.; Minamide, W.; Igarashi, H.; Hinuma, Y.	
#journal		Infect. Immun. (1994) 62:4000-4004	
#title		The gene encoding a new mitogenic factor in a Streptococcus pyogenes strain is a distributed copy in group A streptococci.	
#cross-references		MUID:94341910	
#accession		I39863	
#status		preliminary; translated from GB/EMBL/DBJ	
#molecule_type		DNA	
#residues		1465	#label RES
#cross-references		GB:D12170; NID:9418066; PID:9600252	
SUMMARY		#length 485	#checksum 3782
Quary Match		41.1%	Score 1329; DB 2; Length 485;
Best Local Similarity		39.5%	Pred. No. 1.55e-218;
Matches 172;		Conservative 117;	Mismatches 155; Indels 1; Gaps 1
Db	44	SIDTIGLNTLYNQEVLAANGDKVESFVKESINSNGKFFVVDYVRKNHLDORHQSIFRLD	103
Qy	5	AVNDFILAMNDRKKILLTHGSGESINERFKEBQGNLPDEFVYXERKRKSLSTNTSDIXVTA	64
Db	104	SVANFTTPGAVOLANKAFADNPSSLVAKRAKLNSIDLPGRKRENT-IIVQNPYIGNVA	162
Qy	65	TNDSKLYPGALLVYDETLKKNPTLLAADRAMYISIDLPELASSDFLQVEPSSNSVR	124
Db	163	GAVDIVATVNNKESATHTLPARMQYTESMYSKQOIASALNVAKYLDNSLMDIFENAVA	222
Qy	123	GAVNDLAKKHODYGOVNNVPARXQYERKTLTASHMEQJLKVKGSGFEKTSGLIDENSVHS	184
Db	223	NGEKKVMAIYAKQIFTTVSANLPNNPADYEDKSTYFKLQKGSYNAAPLPFVSNAVYERT	282
Qy	185	SGEQIQIVAKKQIYYTVSYDAKMGKGVDFQDTYVEDLKRGISAEPRLPYISXVAAYGRQ	244
Db	283	TYVYVLETTSSKRYOAKAFALAKKNSVETSQYQDIIEBSTFPNAVILGDAEHKKVYTK	342
Qy	245	QYVTKLETFSSKSEVEEAPFALIKGVKVAPOYEMKQILDNTYKVAVILGGDPSSGARVYTG	304
Db	343	KDFNFRNVIKDNATESRKNPAYISYTSVFLKNNKLAGVNRNTEYVETTSYETSQINL	402
Qy	305	GVMDVEDLIEGSRFPADHPGLDIPYTSFSLRDVNAVTFQNSDYETVETVYARNGDILL	364

Db	403	LDHYGAYVAQFDVSMQGFEPDQNGKELTHHKTGSGKCKTAHYSVIPLPNSKIKIVY	462
QY	365	LDHSGAYVAQYITVWELSTDHQKKEVLTPKAMDNRNGODLTAFHTSTIPLKGVRNLISVK	424
Db	463	ARECTGLAEMWRTI	477
QY	425	IRECTGLAEMWRTV	439
RESULT	10		
ENTRY		547298	#type complete
TITLE		sulysin - Streptococcus suis	
ORGANISM		06-Jan-1995	#formal_name streptococcus suis
DATE		09-Sep-1997	#sequence_revision 06-Jan-1995 #text_change
ACCESSIONS		547298	
REFERENCE		547297	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		547298	
##status		preliminary	
##molecule_type		DNA	
##residues		1-112	#label SEG
##cross-references		EMBL:236908; NID:9535307; PID:9535308	
SUMMARY		#length 112	#molecular-weight 12464 #checksum 9264
Query Match		14.3%;	Score 464; DB 2; Length 112;
Best Local Similarity		50.9%;	Pred. No. 2,28e-57;
Matches		57; Conservative	22; Mismatches 33; Indels 0; Gaps 0
Db	1	PGVPVSTTFFVXXXXPAQXLSNSEIETFTSVHSSALTLKHSAGAYAKYITWGEVSY	60
QY	325	PGLPISYITSELRDNVAATFQNSTDVEYKVAIRNGDLDLHDSAGAYAOYIITWELSY	384
Db	61	NEAGEVWEKXAMKNGVNLISHWSETIXIPENAXNLTHTIOECTGLAEMW	112
QY	385	DHQKEVLLPKAMDNRNGODLTAFHTSTIPLKGVRNLISVKIRECTGLAEMW	436
RESULT	11		
ENTRY		547297	#type complete
TITLE		sulysin - Streptococcus suis	
ORGANISM		06-Jan-1995	#formal_name Streptococcus suis
DATE		09-Sep-1997	#sequence_revision 06-Jan-1995 #text_change
ACCESSIONS		547297	
REFERENCE		547297	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		547297	
##status		preliminary	
##molecule_type		DNA	
##residues		1-96	#label SEG
##cross-references		EMBL:236907; NID:9535305; PID:9535306	
SUMMARY		#length 96	#molecular-weight 10471 #checksum 6662
Query Match		6.6%;	Score 214; DB 2; Length 96;
Best Local Similarity		33.7%;	Pred. No. 3,28e-15;
Matches		30; Conservative	22; Mismatches 36; Indels 0; Gaps 0
Db	1	KQDNQYFOSIYVGPPEILTNGEYIDNPATATGMIENGFRVVLRRKXTTDSADITVI	60
QY	4	KAVNDFILAMNDKKRLTHGQESIEENFEXKRGNOPLPEFVEXERRKRLSTNTSDIXVT	63
Db	61	XAKANISPGALLRAXONLIDNNPLTISI	89
QY	64	ATNDSRLYPGALLVYDETLENNPILLAV	92
RESULT	12		
ENTRY		503974	#type complete
TITLE		amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine	
ALTERNATE_NAMES		monoamine oxidase type A	

```

ORANISM      #journal.name Bos primigenius taurus #common_name cattle
DATE         26-Feb-1990 #sequence_revision 26-Feb-1990 #text_change
ACCESSIONS   #authors
REFERENCE     #journal
             #title
             #cross-references MUID:89246344
             #accession
             #status
             #molecule_type mRNA
             ##residues 1-527 ##label POW
             ##cross-references GB:X15609; NID:9523; PID:9524
KEYWORDS      FAD; flavoprotein; mitochondrion; oxidoreductase
FEATURE       15-43
             #region beta-alpha-beta FAD nucleotide-binding fold\
             modified site S (beta-alpha-PAD)-cysteine (cys) #status
             predicted
SUMMARY       #length 527 #molecular_weight 59800 #checksum 6598

Query Match      3.6%; Score 117; DB 2; Length 527;
Best Local Similarity 19.3%; Pred. No. 1,18e-01;
Matches          29; Conservative 42; Mismatches 73; Indels 6; Gaps 5;

Db 1 MESLKTQDAGQMDVYVIGGIGISGLNAK-LLAHEVNVLYEAREVRGRTYVRNEH 59
    250 LETTSKXSEVAEPALIKGVKAPQTEWKQILDNTEVKAVIIGDPPSGARVYTGKVM 309
    Db 60 V-DYDVAGAVYGPQNILRLSKQLGLEYKVNVRNLVHYKGTTPRGAFPPVMP 118
    QY 310 VEDLIQESRTADHPGLPISTYTSF-LRDNVVAATQNSIDYIVKTYATRNGLDLDS 368
    Db 119 IAYL-DYNNLMRTM--DNMKEIPADAPWE 145
    QY 369 GAYVAQYYITWNELSYDHQKEVLTTPKAMD 398

RESULT 13
ENTRY   S59327 #type complete
TITLE   hypothetical protein YLR135w - yeast (Saccharomyces
        cerevisiae)
ALTERNATE_NAMES  hypothetical protein 13140
ORGANISM         #formal_name Saccharomyces cerevisiae
DATE             29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change
ACCESSIONS       S59327; S64977
REFERENCE
#authors         Delius, H.
#submission      submitted to the EMBL Data Library, June 1995
#description      36.8 kb of S. cerevisiae chromosome XII including ACE2, CK11,
                PBC5, SLS1, PUT1 and tRNA-asp.
                S59327
#molecule_type  DNA
#residues        1-748 ##label DEL
#cross-references EMBL:X91258; NID:995686; PID:995701
#experimental_source strain S288C
REFERENCE
#authors         Delius, H.; Hedling, U.
#submission      submitted to the Protein Sequence Database, May 1996
#accession        S64977
#molecule_type  DNA
#residues        1-748 ##label DEW
#cross-references EMBL:Z73307; NID:g1360550; PID:e245572; PID:g1360551;
                MIPS:YLR135w
#experimental_source strain S288C

GENETICS
#map_position 12R
SUMMARY       #length 748 #molecular_weight 84361 #checksum 5347

```

```

Query Match      3.5%: Score 112: DB 2: Length 748:
Best Local Similarity 18.8%: Pred. No. 4.55e+01:
Matches 33: Conservative 58: Mismatches 71: Indels 14: Gaps 13:

Db 567 SEEFTEFDQFCIADIOLYDSSKISTKSDTOPPTTNSDIIDISAASSIASPEKFEIIM 626
      : : : : : | | | | | : : : : : | | | | | : : : : :
Oy 22 THOGS-TEHFxKXEGNOLDPDEFVxKRRKRSLSTNSDxVYATANDSRKYGAL--LVY 78
      : : : : : | | | | | : : : : : | | | | | : : : : :

Db 627 SOSKELRQSLKTYGLKPMRTKEIILQIOTASOILSTANPDNKGEGVAFN-SK-IEI 684
      : : : : : | | | | | : : : : : | | | | | : : : : :
Oy 79 DELTKENPPTLLADRAPMTYSIDL-PLGASSDSFLQVEDPSSSVRGAVNDLAKMHQD 137
      : : : : : | | | | | : : : : : | | | | | : : : : :

Db 685 FDHLLELEAPPDLERTYFEFPIPLNELLEKLFSAEPFYSQDEITIEKMDVYG 740
      : : : : : | | | | | : : : : : | | | | | : : : : :
Oy 138 YGQVNN-VPAKHQY-EKI-T-AH-SMDLKVK-FGSD-FEKTGNSLDI-DENSVHS 185
      : : : : : | | | | | : : : : : | | | | | : : : : :

RESULT 14
ENTRY F64461 #type complete
TITLE hypothetical protein MJ1295 - Methanococcus jannaschii
ORGANISM #format_name Methanococcus jannaschii
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 10-Oct-1997

ACCESSIONS F64461
REFERENCE A64300
AUTHORS Bul, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; Fitzgerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kierlavage, A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glöck, A.; Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrman, J.L.; Nguyen, D.; Ulteirback, T.R.; Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Rienk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C. Science 1996) 273:1058-1073
#journal Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
#title
#cross-references M01D:96337999
#accession F64461
#status preliminary: nucleic acid sequence not shown; translation not shown

##molecule_type DNA
##residues 1-218 #label BUL
##cross-references GB:U67570; GB:L77117; NID:g1591929; PID:g1591933; TIGR:MJ1295; PDB:g1511304

GENETICS
#map_position REV1243502-1242846
#start_codon TTG
SUMMARY #length 218 #molecular_weight 24666 #checksum 9893

Query Match 3.3%: Score 107: DB 2: Length 218:
Best Local Similarity 26.3%: Pred. No. 1.68e+00:
Matches 21: Conservative 17: Mismatches 40: Indels 2: Gaps 2:

Db 68 IGVVELTHIPKSPKMKAKKADIGAEIVVH-GETVVEPEVEKTKTYVASIEDVDILAH 126
      : : : : : | | | | | : : : : : | | | | | : : : : :
Oy 153 ITASHMEQLKVKFESDPEKFGNSLDIDFNSVHSGEKGIOIVNKKOITYIVYSDAKNRPD 212
      : : : : : | | | | | : : : : : | | | | | : : : : :

Db 127 PGFIDKEFLAEMLKENDIFVE 146
      : : : : : | | | | | : : : : : | | | | | : : : : :
Oy 213 V-FODTVYVEDLKKORGISAE 231
      : : : : : | | | | | : : : : : | | | | | : : : : :

RESULT 15
ENTRY S60947 #type complete
TITLE hypothetical protein YOR220w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein OS050: hypothetical protein YOR50-10
ORGANISM #format_name Saccharomyces cerevisiae
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 05-Jun-1998
ACCESSIONS S60947; S67113; S71722

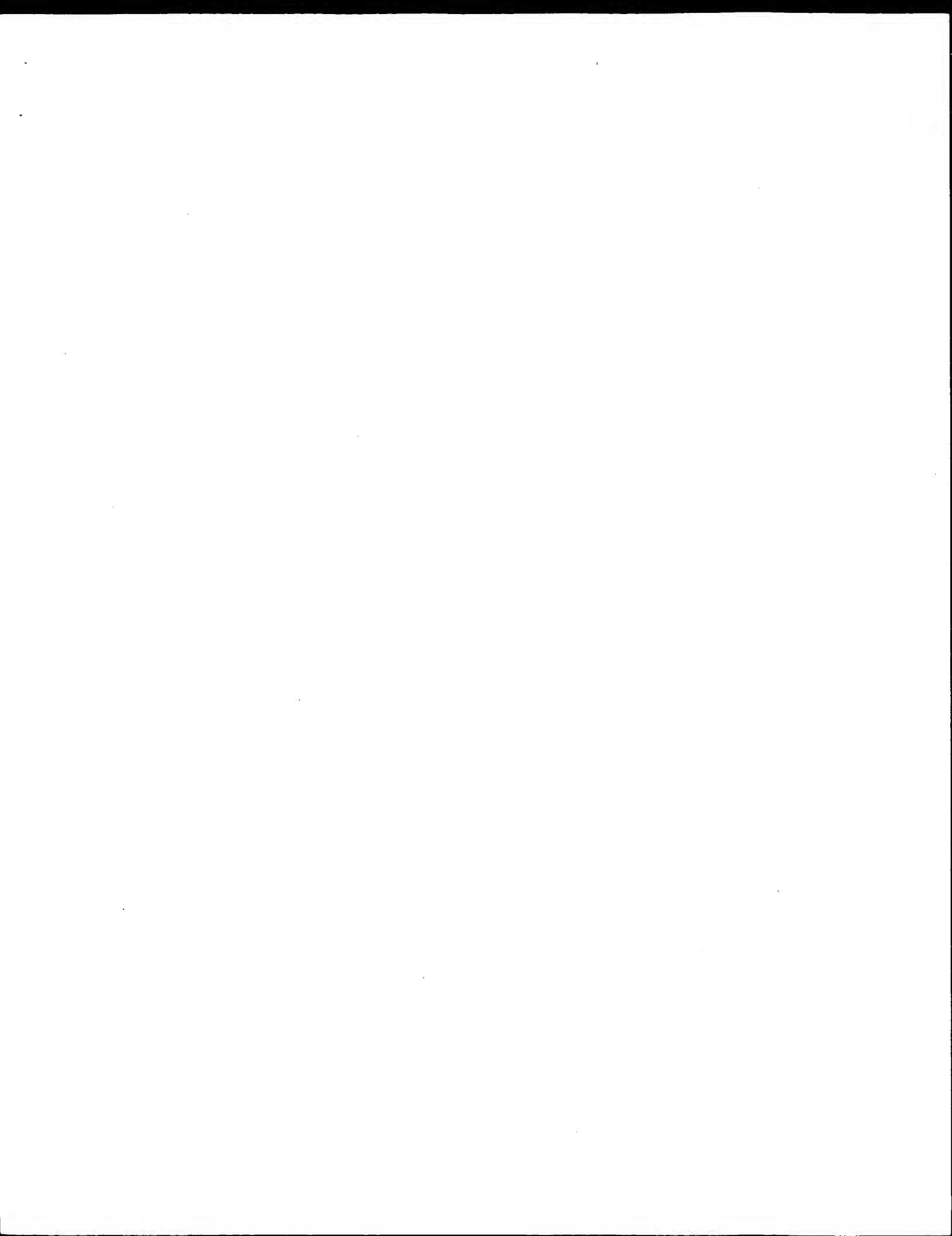
```

```

REFERENCE
#authors          S60938
#submission       Submitted to the EMBL Data Library, October 1995
#description       Sequence and analysis of a 33 kb fragment from the right arm
                   of chromosome XV of the yeast Saccharomyces cerevisiae.
#accession        S60947
#molecule-type   DNA
#residues         1-265 ##label GAL
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
REFERENCE
#authors          Boyer, J.; Fairhead, C.; Gallon, L.; Gallsion, F.; Michaux,
                   G.; Thierry, A.; Dujon, B.
#submission       Submitted to the Protein Sequence Database, July 1996
#accession        S67113
#molecule-type   DNA
#residues         1-265 ##label BOY
#cross-references EMBL:Z75128; NID:g1420509; PID:e252085; PID:g1420510
                   MIPS:YOR220w
#experimental-source strain S288C
REFERENCE
#authors          Gallsion, F.; Dujon, B.
#journal           Yeast (1996) 12:877-885
#title            Sequence and analysis of a 33 kb fragment from the right arm
                   of chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MJD:96437977
#accession        S71722
#status           nucleic acid sequence not shown; translation not shown
#molecule-type   DNA
#residues         1-265 ##label GAW
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
#note             the nucleotide sequence was submitted to the EMBL Data
                   Library, October 1995
GENETICS
#map-position     15R
SUMMARY
#map-position     #length 265 #molecular-weight 29255 #checksum 6028
Query Match          3.3%; Score 106; DB 2; Length 265;
Best Local Similarity 26.7%; Pred. NO. 2.17e+00;
Matches 28; Conservative 26; Mismatches 44; Indels 7; Gaps 7;
Db 85 PAKLFVETSLNNO-HRR-SRSTDDAVSLQNNMLALLEDRHKPLSLNPDGVTGSDS 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 12 AMN-YDKKKLITHGSEIENRFKKGQGLDE-FVYKF-RKKSLSTNSDIXVATINDS 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 SLNKGSSLSPEKSSLESPTMLKISTDSKPFYSQOEPLKLSRSS 187
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 69 RLY-PGALLVVDLTXENNPTL-LAVDRAEMTYSIDLPGLSSDS 111

```

Search completed: Fri Sep 3 13:11:27 1999
Job time : 48 secs.



Matches 460; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Db 1 ANKAVNDPILAMNDKRLTHGEGSIENFIEGNOIPDEFVIERKRSISTNTSDIS 60
 QY 2 ANKAVNFIILAMNDKRLTHGEGSIENFIEGNOIPDEFVIERKRSISTNTSDIX 61
 Db 61 VTANDSRIVPGALVVDPELLNPTLLAVDAPMTYSIDLGLASSDSFLOVEDPSSNS 120
 QY 62 VTANDSRIVPGALVVDPELLNPTLLAVDAPMTYSIDLGLASSDSFLOVEDPSSNS 121
 Db 121 SVRGAVDLAKMHODYGVANNPAPARQYKETAHSHMEOLKVFSGDFEKTGNSLIDEN 180
 QY 122 SVRGAVDLAKMHODYGVANNPAPARQYKETAHSHMEOLKVFSGDFEKTGNSLIDEN 181
 Db 181 SVHSGEQIOIVNFQOITYYVSVDVAVNPGDVFQDITVEDLKQRGISAEPLVYISSVA 240
 QY 182 SVHSGEQIOIVNFQOITYYVSVDVAVNPGDVFQDITVEDLKQRGISAEPLVYISSVA 241
 Db 241 YGRQVYLKLETTSSDSEVAEAFELIKGVAVAPQTEWKQILDNTEVAVILGGDPSSGAR 300
 QY 242 YGRQVYLKLETTSSDSEVAEAFELIKGVAVAPQTEWKQILDNTEVAVILGGDPSSGAR 301
 Db 301 VVTGKVMVEDLIOGSRFTADHPLISYTTSTFLRDNVATFQNSTDYETKVTAYRNG 360
 QY 302 VVTGKVMVEDLIOGSRFTADHPLISYTTSTFLRDNVATFQNSTDYETKVTAYRNG 361
 Db 361 DLLIDHSGAVYAOYITWDELSDYHOGKEVLTPKAMRNGODLTAHFTTISPLKGNVRL 420
 QY 362 DLLIDHSGAVYAOYITWDELSDYHOGKEVLTPKAMRNGODLTAHFTTISPLKGNVRL 421
 Db 421 SVKIRETGIAEMWRTVYEKTDLPYRKRTISMGTTLPQVEDKEND 470
 QY 422 SVKIRETGIAEMWRTVYEKTDLPYRKRTISMGTTLPQVEDKEND 471

RESULT 2
 ID TACY CLOPE STANDARD; PRT; 500 AA.
 AC P19995;
 DT 01-FEB-1991 (REL. 17, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE PERRINGOLYSIN O GENE OF CLOSTRIDIUM PERRINGENS."
 GN PRO OR PRO OR PROA.
 OS CLOSTRIDIUM PERRINGENS.
 OG PLASMID PR11B.
 OC BACTERIA; FRMUTURES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
 CC CLOSTRIDIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 9109951.
 RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
 RT "An upstream regulatory sequence stimulates expression of the
 RT perringolysin O gene of Clostridium perringens.";
 RL INFECT. IMMUN. 59:137-142(1991).
 RN [2]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 29-44.
 RX STRAIN=ATCC 13124;
 RC MEDLINE; 89032623.
 RA TWETEN R.K.;
 RT "Nucleotide sequence of the gene for perringolysin O (theta-toxin)
 RT from Clostridium perringens: significant homology with the genes for
 RT streptolysin O and pneumolysin.";
 RL INFECT. IMMUN. 56:323-324(1988).
 RN [3]
 RP SEQUENCE OF 29-45 AND 305-312.
 RC STRAIN=PB6K;
 RX MEDLINE; 87076517.
 RA OHNO-IMASHITA Y., IMAOTO M., MITSUI K., KAMASAKI H., ANDO S.;
 RT "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
 RT from Clostridium perringens.";
 RL BIOCHEMISTRY 25:6048-6053(1986).
 RN [4]
 RP SEQUENCE OF 492-500 FROM N.A.

RC STRAIN=NCCTC 8237;
 RX MEDLINE; 96123363.
 RA SHIMIZU T., KOBAYASHI T., BA-THEIN W., OHTANI K., HAYASHI H.;
 RT "Sequence analysis of flanking regions of the pfoA gene of
 RT Clostridium perringens: beta-galactosidase gene (pbg) is located in
 RT the 3'-flanking region.";
 RL MICROBIOL. IMMUNOL. 39:677-686(1995).
 RN [5]
 RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
 RX MEDLINE; 88004463.
 RA IMAOTO M., OHNO-IMASHITA Y., ANDO S.;
 RT "Role of the essential thiol group in the thiol-activated cytolysin
 RT from Clostridium perringens.";
 RL EUR. J. BIOCHEM. 167:425-430(1987).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE; 97113505.
 RA FEIL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Crystallization and preliminary x-ray analysis of a thiol-activated
 RT cytolysin.";
 RL FEBS LETT. 397:290-292(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE; 97325744.
 RA ROSSJOHN J., FEIL S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
 RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
 RT model of its membrane form.";
 RL CELL 89:685-692(1997).
 CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 CC EUKARYOTIC CELL MEMBRANES.
 CC -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M81080; G144886; -;
 DR EMBL; M36704; G144884; -;
 DR EMBL; D49537; G1502275; -;
 DR PIR; B43577; B43577;
 DR PDB; 1PFO; 05-AUG-98.
 DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
 DR PFM; PF01289; THIOL_CYTOLYSIN; 1.
 KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
 FT SIGNAL 1 28
 FT CHAIN 29 500 PERRINGOLYSIN O.
 FT ACT SITE 459 459 BINDING TO CHOLESTEROL.
 FT CONFLICT 30 30 D -> K (IN REF. 3).
 FT CONFLICT 34 34 K -> I (IN REF. 3).
 FT CONFLICT 126 128 RKP -> EA (IN REF. 2).
 SQ SEQUENCE 500 AA; 55799 MW; 17FC32BC CRC32;

Query Match 49.5%; Score 1600; DB 1; Length 500;
 Best Local Similarity 46.5%; Pred. No. 0.00e+00;
 Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;

Db 35 NOSIDGISTSLYNEVLAISNGDKIESFYKKGKAGKRIYEROKRSLTSPVDISI 94
 QY 3 NKAVDNFIILAMNDKRLTHGEGSIENRFPKKEGNOIPDEFVIERKRSISTNTSDIX 62
 Db 95 IUSVNDRIYPPGALADAFVFNRPITIMVAKKININIDLPGLKGENS-1KYDDPIYTGK 153
 QY 63 TATNSRSLYPPGALVVDPELLNPTLLAVDAPMTYSIDLGLASSDSFLOVEDPSSNS 122
 Db 154 VSGAIDELVSKWNEKYSSTHTLPARTOYSESMVYSKSGISSALVWNAKVLENSIGVDFNA 213

[illegible]

```

Oy      130 LIAKMHODYGVNNVPAKXOYEKITAHSMDEQLAKFCFSDEPKTNSLDIDFNSVHSGEKQ 189
Db      223 VMVAAYKOIFETVSAGLEPNPNSDIFEDSVTFPAELARKGVSNEAPLMAVNVAIRITIVK 282
Oy      190 IOIWMKKOIYYTVSDAVKANGDFODTVEJVEDLKORISAEPLVYIXSVAVGROYLYK 249
Db      283 LETTSKSNDOVTARFKLLINPNSIQASGGYKOIYENSSFTAVVLGGDDQTHNOVTKDFNV 342
Oy      250 LETTSKSXEVPAAEPALIKGVKVAPOTEMKOILNTEYKAVALLGDSGSGARVVTKGYDM 309
Db      343 IOSVLIKDAWESSKNKPAPISITYTSVFTEKDNSIAAHHNTETIETRTYESKGRIKLHDGS 402
Oy      310 VEDLIQESRRTADHPGLPISTYSFLDNVANVAFFQNSTDYVERKTVAYVRNGDLLDHSG 369
Db      403 AVYAOFEEYWDSEFDADGOEITVTRKSKDWGMWRBRSASFTEIEPLPPAKAIRIPARECT 462
Oy      370 AYVAQYITTWMLSDYDHOGKEVLTPEKADRNGODTLAHTFTSIPLKGNVMLSVKIRECT 429
Db      463 GIAMEMWRTVVDENVVPLASDIANSIWCTILXP 495
Oy      430 GIAMEMWRTVYEKTDLPVLRKRRTISIMWTILXP 462

RESULT 7
AC TACY_STREQ STANDARD: PRT: 574 AA.
AC Q54114.
Dt 15-JUL-1998 (REL. 36, CREATED)
Dt 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
Dt 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
De STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
Gn SLO.
Os STREPTOCOCCUS EQUISTMILIS.
Oc BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
Cc STREPTOCOCCUS.
Cn [1]
Rp SEQUENCE FROM N.A.
Rc STRAIN-SIBD-1.
Rx MEDLINE: 95102113.
Rx OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
Ra IGARASHI H., YUTSUO T.;
Rt "Cloning and sequencing the streptolysin O genes of group C and group
Rt G streptococci."
Rl DNA SEQ. 4:325-328(1994).
Cc -I- FUNCTION: SOLUBLELY-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
Cc CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVAED BY OXIDATION.
Cc CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
Cc EUKARYOTIC CELL MEMBRANES.
Cc -I- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as their content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL: D16824; G498301; -.
Cc PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
Cc PRAM: PRO1289; Thiol_Cytolysin; 1.
Cc HSP: P19995; IPFO.
Cc KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
Cc FT SIGNAL 1 36 BY SIMILARITY.
Cc FT CHAIN 37 574 STREPTOLYSIN O.
Cc FT ACT SITE 533 533 BINDING TO CHOLESTEROL (BY SIMILARITY).
Cc SQ SEQUENCE 574 AA; 63991 MW; 9BBP31B4 CRC32;

Query Match 43.68; Score 1409; DB 1; Length 574;
Best Local Similarity 41.6%; Pred. No. 4.24e-271;
Matches 190; Conservative 106; Mismatches 160; Indels 1; Gaps 1;

Db 112 INDKIYSLANNELEVAKNGETTELVNPRKGVKAKFIYIERKKKNIMTTTPVDISIIDS 171
::| | ::||| :|: ||:||| ||| :| | | |||||:::| | : :

```

```

QY 6 VNDPILAMNDKKLLTHOGESIENRFXKEGNQLPDEFVYXERKRKSLSTNTSDIXVAT 65
DB 172 VDRTPAALOLANKGFENKPDPAVYTKRNPQKHIDLPGMGDKAT-VEVNDPTANST 230
QY 66 NDRSLYPGALLVYDELXENNPTLLAVDAPMTYSIDLPGGLASSPSFLQEDPSSSVRG 125
DB 231 AIDNLVNMHDNYSNGNTLPARTQYTESNYSKSOIEALNYSKILDTGLIDFKSISK 290
QY 126 AVNDLLAKHODYGGVANNPARXQYEKITAHSMEOLKVFSGDFEKTGNSLIDENSVHS 185
DB 291 GEKKVIAAOKOIFYVSANLPNNADVPDKSVTLKEIQRKVSNEAPPLVSNVAGRT 350
QY 186 GEKOIOIVXKKOITYTVSDAVKNPQVDFQDVTVEDLKORGISAEPLVYISXVAYGRQ 245
DB 351 VEVKLETSSKSDNDEAFAALKGTVDKNGKYSIDLENSFTAVLGGDAEHNKVVTK 410
QY 246 VYLKLETTSSXSEVEAFALIKGVKVAPOQEMKQILDNTIEKAVIILGDPSSGARVYTG 305
DB 411 DEDVIRNVYKDNATSRKRPAPVITSYVFLKNNKIAGVNNREVEVETTSYTGSKINTL 470
QY 306 KYDMVEDLQEGSRFTADHPGLPISTYSFLDNVAVAFQNSTDYVETKVAAYRNGDLL 365
DB 471 SHOGAYVAQYELIMPEINVDKGEVITKRRMDNNWYSKTSFSPVITPLGANSRIRIMA 530
QY 366 DHSAGVVAQYITTMWELSYDHOGKEVLPKAMDNRGQDLTAHFTTSLPKGNVRLSVKI 425
DB 531 RECTGLAMEMWRKVIDERDVYLSKEINYNISGSTLSP 567
QY 426 RECTGLAMEMWRVYEKTDLPVLRKRRTISIMGTILYP 462

RESULT 8
ID TACY_STRPY STANDARD: PRT: 571 AA.
AC P21131;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC PLASMID PMK157.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
RN STREPTOCOCCUS.
RP [1]
RA SEQUENCE FROM N.A.
RA MEDLINE: 88057628.
RA KEROE M.A., MILLER L., WALKER J.A., BOULNOIS G.J.;
R1 "Nucleotide sequence of the streptolysin O (SLO) gene: structural
homologies between SLO and other membrane-damaging, thiol-activated
toxins."
RT INECT. IMMUN. 55:3228-3232(1987).
CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M18638; G153811; .
CC PIR: A43507; A43507.
CC PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
CC PRAM: PFO1289; Thiol_cytolysin; 1.
CC HSSP: P19995; 1PFO.
CC TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID.
FT SIGNAL 1 33
CHAIN 34 571 STREPTOLYSIN O.

```

```

FT ACT SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA: 63638 MW: 33124534 CRC32;
Query Match 43.5%; Score 1408; DB 1; Length 571;
Best Local Similarity 41.8%; Pred. No. 7 07e-271;
Matches 191; Conservative 104; Mismatches 161; Indels 1; Gaps 1;

DB 109 INKITSLNNELEVAKGETIENRVEPEGVKAKDFIVERRKKNINTTPVDISIDS 168
QY 6 VNDPILAMNDKKLLTHOGESIENRFXKEGNQLPDEFVYXERKRKSLSTNTSDIXVAT 65
DB 169 VDRTPAALOLANKGFENKPDPAVYTKRNPQKHIDLPGMGDKAT-VEVNDPTANST 227
QY 66 NDRSLYPGALLVYDELXENNPTLLAVDAPMTYSIDLPGGLASSPSFLQEDPSSSVRG 125
DB 228 AIDNLVNMHDNYSNGNTLPARTQYTESNYSKSOIEALNYSKILDTGLIDFKSISK 287
QY 126 AVNDLLAKHODYGGVANNPARXQYEKITAHSMEOLKVFSGDFEKTGNSLIDENSVHS 185
DB 288 GEKKVIAAOKOIFYVSANLPNNADVPDKSVTLKEIQRKVSNEAPPLVSNVAGRT 347
QY 186 GEKOIOIVXKKOITYTVSDAVKNPQVDFQDVTVEDLKORGISAEPLVYISXVAYGRQ 245
DB 348 VEVKLETSSKSDNDEAFAALKGTVDKNGKYSIDLENSFTAVLGGDAEHNKVVTK 407
QY 246 VYLKLETTSSXSEVEAFALIKGVKVAPOQEMKQILDNTIEKAVIILGDPSSGARVYTG 305
DB 408 DEDVIRNVYKDNATSRKRPAPVITSYVFLKNNKIAGVNNREVEVETTSYTGSKINTL 467
QY 306 KYDMVEDLQEGSRFTADHPGLPISTYSFLDNVAVAFQNSTDYVETKVAAYRNGDLL 365
DB 468 SHOGAYVAQYELIMPEINVDKGEVITKRRMDNNWYSKTSFSPVITPLGANSRIRIMA 527
QY 366 DHSAGVVAQYITTMWELSYDHOGKEVLPKAMDNRGQDLTAHFTTSLPKGNVRLSVKI 425
DB 528 RECTGLAMEMWRKVIDERDVYLSKEINYNISGSTLSP 564
QY 426 RECTGLAMEMWRVYEKTDLPVLRKRRTISIMGTILYP 462

RESULT 9
ID TACY_STRCB STANDARD: PRT: 574 AA.
AC Q53957;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS CANIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
RN STREPTOCOCCUS.
RP [1]
RA SEQUENCE FROM N.A.
RA MEDLINE: 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
R1 "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
CC -!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```



```

CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEAMINATION OF BIOGENIC AND
CC XENOBOTIC AMINES AND HAS IMPORTANT FUNCTIONS IN THE METABOLISM OF
CC NEURORACTIVE AND VASOACTIVE AMINES IN THE CENTRAL NERVOUS SYSTEM
CC AND PERIPHERAL TISSUES. MAO-A PREFERENTIALLY OXIDIZES BIOGENIC
CC AMINES SUCH AS 5-HYDROXYTRYPTAMINE (5-HT), NORPINEPHRINE AND
CC EPINEPHRINE.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: MONOMER, HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15609; G524; -
CC DR EMBL; X15609; G525; ALT_INIT.
CC DR PIR; S03974; S03974.
CC KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; TRANSMEMBRANE; MITOCHONDRION;
CC NEUROTRANSMITTER DEGRADATION.
CC FT NP_BIND 14 70 FAD (ADP PART) (POTENTIAL).
CC FT BINDING 406 406 FAD (BY SIMILARITY).
CC FT TRANSMEM 498 518 POTENTIAL.
CC SQ SEQUENCE 527 AA; 59800 MW; 724E4396 CRC32;

```

Query Match 3.68; Score 117; DB 1; Length 527;
 Best Local Similarity 19.38; Pred. No. 2,05e-02;
 Matches 29; Conservative 42; Mismatches 73; Indels 6; Gaps 5;

```

Db 1 MESLQKTS DAGQMFVYVVGIGGIGSLGSAK-LLAEHEVVLLEARENGGTYVRNEH 59
QY 250 LETTSXSEVEAFELINGVAVAPOTEKQILDITEVAAILGDPSSGARVTVGKVD 309
Db 60 V-DYDVGGAYVGPIONRLRLSKQLGLETYYVNVNERLVHYKGTYPFGCAFPPVNP 118
QY 310 VEDLIDGSRFTADHGLPISTYTSF-LRDNVAVAFQNSTDYVEIKVAYNRGDLIDHS 368
Db 119 IAYL-DYNNLMRTM--DNMGKEIPADAPWE 145
QY 369 GAVVAQVYITWNLSTYDHQKEVLTPKAWD 398

```

RESULT 12
 ID MPEE YEAST STANDARD: PRT: 462 AA.
 AC P10507;
 DT 01-JUL-1988 (REL. 11, CREATED)
 DT 01-JUL-1988 (REL. 11, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT PRECURSOR
 DE (EC 3.4.24.64) (BETA-MP) (PEP)
 GN MASI OR MTF1 OR YLR163C OR YL9632.10.
 OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACHAROMYCETALES;
 OC SACHAROMYCETACEAE; SACHAROMYCES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 88312592.
 RA WITTE C., JENSEN R.E., YAFFE M.P., SCHATZ G.;
 RT "MAS1, a gene essential for yeast mitochondrial assembly, encodes a
 RL subunit of the mitochondrial processing protease.";
 RL EMBL J. 7:1439-1447(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN: S288C / AB972;
 RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,

```

RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
RA KUCABA T., HALLSWORTH K., HAWKINS J., HILLIER L., JIER M.,
RA JOHNSON D., JOHNSON L., LANGSTON Y., LATREILLE P., LE T.,
RA JARDIS E., HENNES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
RA RIRKEN L., RILES L., TAICH A., TREVASKIS E., VIGNATI D.,
RA WILCOX L., WOHLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;
RA SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 21-32.
RX MEDLINE; 91177897.
RA YANG M., GELI V., OPLIGER W., SUDA K., JAMES P., SCHATZ G.;
RT "The MAS-encoded processing protease of yeast mitochondria.
RT Interaction of the purified enzyme with signal peptides and a
RT purified precursor protein."
RL J. BIOL. CHEM. 266:6416-6423(1991)
CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF N-TERMINAL TRANSIT PEPTIDES FROM
CC PRECURSOR PROTEINS IMPORTED INTO THE MITOCHONDRION, TYPICALLY WITH
CC ARG IN POSITION P2.
CC -1- COFACTOR: REQUIRES DIVALENT CATIONS FOR ACTIVITY.
CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X07649; G3887; -
CC DR EMBL; U51921; G1234852; -
CC DR PIR; S00552; S00552.
CC DR PIR; A38734; A38734.
CC DR SGD; L0001026; MASI.
CC DR PROSITE; PS00143; INSULINASE: 1.
CC DR PFAM; PF00675; Insulinase: 1.
CC KW HYDROLASE; METALLOPROTEASE; ZINC; MITOCHONDRION; TRANSIT PEPTIDE.
CC FT TRANSIT 1 20 MITOCHONDRION.
CC FT CHAIN 21 462 MITOCHONDRIAL PROCESSING PROTEASE
CC FT METAL 70 70 BETA SUBUNIT.
CC FT ACT SITE 73 73 ZINC (BY SIMILARITY).
CC FT METAL 74 74 ZINC (BY SIMILARITY).
CC FT METAL 150 150 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 462 AA; 51083 MW; 438D1777 CRC32;

```

Query Match 3.3%; Score 107; DB 1; Length 462;
 Best Local Similarity 22.38; Pred. No. 4.15e-01;
 Matches 33; Conservative 44; Mismatches 61; Indels 10; Gaps 9;

```

Db 68 TAFLEHLAFL-GTQ-NRSQGIETLEIENI-GSHLNAYTSRENTYVYAKSLQEDIPKAVD 124
QY 154 TASHMQLVKFKGSDPEKGNLSDIDFNSHSGEKOIVNKKOYIVYV-S-DAKNGD 212
Db 125 IISDLTKSVLNSAIERRDYIIESEVYDKMAYEVFDHLEHTTYDQPLGRIIICPT 184
QY 213 VFQDVTVEEDLKQGISARPLVYI-Sx-VA--YGRQYVLYL-ETTSXSEVEAFELI 267
Db 185 KNKISITRDLKDYIFKNKYGDMVLAG 212
QY 268 KGKVAAPQTEKMO-IIDNTEVKAIVILGG 294

```

RESULT 13
 ID NOPS YEAST STANDARD: PRT: 511 AA.
 AC Q12499;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DR EMBL: X04505; G47471; .
DR PIR: S07375; S07375.
DR PRAM: PF00669; Flagellin_N; 1.
DR PRAM: PF00700; Flagellin_C; 1.
KW FLAGELLA.
FT INIT_MET
SQ SEQUENCE 492 AA; 51285 MW; E0DB5DB4 CRC32;

Query Match 3.28; Score 104; DB 1; Length 492;
Best Local Similarity 20.38; Pred. No. 9.84e-01;
Matches 43; Conservative 64; Mismatches 91; Indels 14; Gaps 11;

Db 94 LAVOSANSTNSQSDLDLSIQAEITQRLNEIDRVSGOTQFNGVKVLAODNLTITQGAN-D- 151
QY | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
113 LQVEDPENSSVGRGVNDLLAKMHODYGVNNVPAKQYEEKITAHSMQ-LKVEGSDFEK 171
Db 152 -GETIDIDLKQINSQTLGLDTLVQOKY-KVSDPAATVIG-YTDSATAID-KSTFAASA 206
QY | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
172 TGNSLDIDFNSVHSGEKOIIVNKKQIYTVSDAVKNPGDVFQDIYVEDLKQRCISAE 231
Db 207 TLLGGTPAIT-G-D--LKPDDTGKYYADVSGTTAKDCYEVTVAAADGKVTLTGPTGPI 262
QY | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
232 RPLVYISXVAYGRQVYIKLETTSSKSEVEAFALIKGV-KVAPQTEWKQIIDNTEVKAV 290
Db 263 T-AGFPSTATKDVKQTOQENADLTFAKAAITA 293
QY | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
291 ILGGDPSSGARVYTGKVDWVEDLIQEGSRFTA 322

Search completed: Fri Sep 3 13:08:42 1999
Job time : 55 secs.


```

RT "Listeria monocytogenes exists in at least three evolutionary lines:
RT Listeria from flagellin, invasive associated protein and
RT listeriolysin O genes.";
RL MICROBIOLOGY 141:2053-2061(1995).
DR EMBL; X85855; G940601; -.
FT NON_TER 1 1
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5368 MW; 7080399E CRC32;

Query Match 3.8%; Score 124; DB 2; Length 50;
Best Local Similarity 34.7%; Pred. No. 1,83e-03;
Matches 17; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

Db 2 NSSFKAVIYGSGAKDQVOIDGNLGDLRDLKKGATFNRRPGVIAT 50
|: |||| |:: :::: :::: |::: |::: |::: |::: |::: |
Qy 284 NTEKAVILGGDPSSGARVYTKVDWEDLLQESRFTADHPGLISYT 332

RESULT 5 PRELIMINARY; PRT; 50 AA.
ID Q48773 AC Q48773:
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O (FRAGMENT).
GN HLY.
OS LISTERIA MONOCYTOGENES.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=206.1.1, TYPE 2A;
RX MEDLINE; 96118685.
RA RASUUSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;
RT "Listeria monocytogenes exists in at least three evolutionary lines:
RT evidence from flagellin, invasive associated protein and
RT listeriolysin O genes.";
RL MICROBIOLOGY 141:2053-2061(1995).
DR EMBL; X85935; G940637; -.
FT NON_TER 1 1
FT VARIANT 13 13 F -> Y.
FT NON_TER 50 50
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5429 MW; A4B59721 CRC32;

Query Match 3.6%; Score 118; DB 2; Length 50;
Best Local Similarity 34.7%; Pred. No. 1,29e-02;
Matches 17; Conservative 14; Mismatches 18; Indels 0; Gaps 0;

Db 2 NSSFKAVIYGSGAKDEVCIIDGNIIDLRIILKKAGTFNRPGVIAT 50
|: |||| |:: :::: :::: |::: |::: |::: |::: |::: |
Qy 284 NTEKAVILGGDPSSGARVYTKVDWEDLLQESRFTADHPGLISYT 332

RESULT 6 PRELIMINARY; PRT; 450 AA.
ID Q68518 AC Q68518:
DT 01-AUG-1998 (TREMBREL. 07, CREATED)
DT 01-AUG-1998 (TREMBREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE HYPOHETICAL 49.7 KD PROTEIN.
CS MYXOCOCCUS XANTHUS.
~ BACTERIA: PROTEOBACTERIA; DELTA SUBDIVISION; MYXOBACTERIA;
OC MYXOCOCCALES; CYSTOBACTERINAE; MYXOCOCCACEAE; MYXOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DZFL;
KA MEDLINE; 96347136.
RA TRUDEAU K.G., WARD M.J., ZUSMAN D.R.;
RT "Identification and characterization of FrzZ, a novel response
RT regulation necessary for swarming and fruiting-body formation in
RT Myxococcus xanthus.";
RL MOL. MICROBIOL. 20:645-655(1996).
RN [2]
```

```

RP SEQUENCE FROM N.A.
RA STRAIN-D2F1.
RC WARD M.J., ZUSMAN D.R.:
RL SUBMITTED (FEB-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RM EMBL, AF049107; G2947295; -.
RS HYPOTHETICAL PROTEIN.
SQ SEQUENCE 450 AA; 49686 MW; E3A6FC67 CRC32;

Query Match 3.5%; Score 113; DB 2; Length 450;
Best Local Similarity 26.8%; Pred. No. 6,29e-02;
Matches 30; Conservative 26; Mismatches 49; Indels 7; Gaps 6

Db 9 PLVATVSDMTAPKLI-LFADPKGRVMEHPYILATLRSGEELVPPQD-KPLDPLSTG-RLV 65
||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 233 PLVYI-SXVAGRGROYLTLETTSSXSEVEAFELILIGKVAAPOTEMKQI-LDNTVEKAV 290
HLPGLPLPGLPHEPGELETLVREMGVGTFFPNVAVGALLPPGYRTFLPGEV 117
Db 66 HLPGLPLPGLPHEPGELETLVREMGVGTFFPNVAVGALLPPGYRTFLPGEV 117
Oy 291 ILGGDPSGARVYTGKVDMEEDLIQESGRTYADPG--LPISITTSFLRDNV 340

RESULT 7
ID 012098 PRELIMINARY; PRT; 748 AA.
AC 012098;
DT 01-NOV-1996 (TREMBLREL. 01. CREATED)
DT 01-NOV-1996 (TREMBLREL. 01. LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01. LAST ANNOTATION UPDATE)
DE CHROMOSOME XII READING FRAME ORF YLR135W.
L19606.6 OR L3140.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RA DELTUS H., HEBLING U.:
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA MIRS.
RL SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WILCOX L.:
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HALLSMOOTH K., HANKINS J., HILLIER L., JIER M., JOHNSON D.,
RA JOHNSTON L., LANGSTON Y., LATRILLE P., LE T., MARDIS E., MENEZES S.,
RA MILLER N., NHAN M., PAULEY A., PELUSO D., RIFKEN L., RILES L.,
RA TALIG A., TREVASKIS E., VIGANTI D., WILCOX L., WOLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.:
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WATERSTON R.:
RL SUBMITTED (APR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RA DELTUS H.:
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL, Z73307; E245572; -.
DR EMBL, U53881; G1256901; -.
DR EMBL, X91258; G995701; -.
SQ SEQUENCE 748 AA; 84361 MW; 4900A1F6 CRC32;

Query Match 3.5%; Score 112; DB 3; Length 748;
Best Local Similarity 18.8%; Pred. No. 8.58e-02;

```

Matches 33; Conservative 58; Mismatches 71; Indels 14; Gaps 13;

LD 567 SEEEFEFDQCIADIQVSSKISTKSTONPTTNDIIDTSAASIAPEKCEIMM 626
 QY 22 THOGES-IENRFKKEGNQLPDEFVAKERRKSLSTNTSDIXYATINDSLYALGAL-LVV 78
 527 SOSMKELROSILTVGLKPKRTKEVLIQISQISLSTANPNKKEGCVANF-SK-IEI 684
 QY 79 DETIKENNTTLAVDRAPMTYSIDL-PGLASSDSFLQYEDPSPNSVSKRAYNDLAKMHOD 137
 DB 685 FPHLELIPAPDFLEIRITFEPIPIINLEIKLFSAEFPVSQIDMTREMAVVG 740
 QY 138 YQGVN-VBARQY-EKI-T-AH-SMEQLKVA-FGSD-FERTGNSLDI-DENSVHS 185

RESULT 8
 ID Q58691 PRELIMINARY; PRT; 218 AA.
 AC 058691.
 DT 01-JUN-1998 (TREMELREL, 06, CREATED)
 DT 01-JUN-1998 (TREMELREL, 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMELREL, 06, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN M01295.
 GN M01295.
 OS METHANOCOCCUS JANNASCHII.
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96337999.
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
 RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
 RA SCOTT J.L., GEOHAGAN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
 RA COTTON M.D., ROBERTS C.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
 RA KLEIN H.-P., FRASER C.M., SMITH H.O., WOESSE C.R., VENTER J.C.,
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RT Jannaschii.";
 RL SCIENCE 273:1058-1073(1996).
 CC -1 SIMILARITY: STRONG, TO A. FULGIDUS AF1550 AND
 CC M. THERMOTUTROPHICUM MTH576.
 CC EMBL; 067570; G1591933; -.
 DR HYPOTHETICAL PROTEIN.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 218 AA; 24666 MW; CFC99AB9 CRC32.

Query Match 3.3%; Score 107; DB 1; Length 218;
 Best Local Similarity 26.3%; Pred. No. 3.95e-01;
 Matches 21; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

DB 68 IGVVELTHIPKRSIPKAKKADLGAELIYVH-GETVPEVEPEKNTYASISHDVILAH 126
 QY 153 ITASMSQMLKKEGSDPEKNGNSLDDIFNSVHSEKQIOLVKNKQIYTVADAVKPNPD 212
 DB 127 PGFIDKETAENLKENDIVE 146
 QY 213 V-FODTVTEDLKORGISAE 231

RESULT 9
 ID Q12044 PRELIMINARY; PRT; 265 AA.
 AC Q12044;
 DT 01-NOV-1996 (TREMELREL, 01, CREATED)
 DT 01-NOV-1996 (TREMELREL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMELREL, 01, LAST ANNOTATION UPDATE)
 DE CHROMOSOME XV READING FRAME ORF YOR220W.
 OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCAROMYCETALES;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
 RA THIERRY A., DUON B.;

RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS.
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FY1679;
 RA GALLISSON F., DUON B.;
 RL SUBMITTED (OCT-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; 275128; E252085; -.
 DR EMBL; X92441; G1050772; -.
 SQ SEQUENCE 265 AA; 29255 MW; 2D5D78E CRC32.

Query Match 3.3%; Score 106; DB 3; Length 265;
 Best Local Similarity 26.7%; Pred. No. 5.33e-01;
 Matches 28; Conservative 26; Mismatches 44; Indels 7; Gaps 7;

DB 85 PMKLFVETSLNNO-HPR-SRSTDAVSLODNNLALLDHRRKPLSTINTDGVYGVDS 142
 QY 12 AMN-YDKKLLTHOGESLENRFKKEGNQLPDE-FVYKE-RKRSLSTSDIXYATINDS 68

DB 143 SLKNGSSSLSPKSSLESPYMKLSTDSKPSYQEPKLSRSS 187
 QY 69 RLX-PGALLVDETLXENPPL-LAVDRAPMTYSIDLPGIASSDS 111

RESULT 10
 ID 028947 PRELIMINARY; PRT; 501 AA.
 AC 028947;
 DT 01-JAN-1998 (TREMELREL, 05, CREATED)
 DT 01-JAN-1998 (TREMELREL, 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMELREL, 07, LAST ANNOTATION UPDATE)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN AF1322.
 OS ARCHAEoglobus FULGIDUS.
 OC ARCHAEA; EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobaceae;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE; 98049343.
 RA KLEIN H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
 RA PETERSON S., REICH C.I., MCNEIL L.K., BADGER J.H., GLODEK A., ZHOU L.,
 RA OVERBECK R., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,
 RA COTTON M.D., SPRIGGS T., ARTACH P., KAINE B.P., SYKES S.M.,
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESSE C.R.,
 RA VENTER J.C.;
 RT "The complete genome sequence of the hyperthermophilic,
 RT sulphate-reducing archaeon Archaeoglobus fulgidus.";
 RL NATURE 390:364-370(1997).
 DR EMBL; AE001012; G2649254; -.
 DR TIGR; AF1322; -.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 501 AA; 55698 MW; 5EE0081E CRC32.

Query Match 3.3%; Score 107; DB 1; Length 501;
 Best Local Similarity 27.5%; Pred. No. 3.95e-01;
 Matches 22; Conservative 24; Mismatches 31; Indels 3; Gaps 3;

DB 217 LVANDGVEYGRIVE-VNGVYSSVDYKAVQNNELAEIKIVNDEIRILISYPAVMGY 275
 QY 250 LETSXSXEYBAFAELIKYKVAPOTEMKQILDNTEKAV-ITGGDPSSGARVYTKVD 308

DB 276 KVIGLYTENGEKPPALAGI 295
 QY 309 MVEDL-IOEGSRFTADHPTL 327

Query	246	YLKLETTSSXSEVEAAFEALING	269
RESULT	13		
ID	087083	PRELIMINARY;	PRT; 1361 AA.
AC	087083;		
DT	01-NOV-1998 (TREMBL.REL. 08, CREATED)		
DT	01-NOV-1998 (TREMBL.REL. 08, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBL.REL. 08, LAST ANNOTATION UPDATE)		
DE	S-LAYER PROTEIN.		
OS	CAMPYLOBACTER RECTUS.		
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;		
OC	CAMPYLOBACTER.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 33238;		
RX	MEDLINE; 98442421.		
RA	MIRAMOTO Y.; MAEDA H.; KITANAKA M.; KOKEGUCHI S.; TAKASHIBA S.,		
RT	MURAYAMA Y.;		
RT	"The S-layer protein from Campylobacter rectus: sequence		
RT	determination and function of the recombinant protein.";		
RL	FEMS MICROBIOL. LETT. 166:275-281(1998).		
DR	EMBL; AB001876; D1034498; -		
KW	S-LAYER.		
SO	SEQUENCE	1361 AA; 144904 MW; 55A9D736 CRC32;	
Query Match	3.3%;	Score 106; DB 2; Length 1361;	
Best Local Similarity	20.4%;	Pred. No. 5.33e-01;	
Matches	38; Conservative	49; Mismatches 89; Indels 10; Gaps 10;	
Db	178	VDMSNPHPKATLEITNDNATANFNAPMK-HNFGCDRIITLQSSDKLTGDIYSRHD	236
Qy	115	VEDPSSNVSRQAVNDLAKMHQDYGVNNPAPKQYERKITAHSMELQKVKFGSDFEKTG	173
Db	237	NLTWVEFGQNADEGDPTRPTLTNTIONINIEVTGVTNLDL-RDSNDYEKINIHRTK	295
Qy	174	NSLDIDFNSVSGE-KQI-QIVNAKQI-YTVASDAKNEDVFQDIYVTEDLKQKQISA	230
Db	296	EAGKKEFVESIG-QKLVGMRLANYAK-KDIDVKEH-KKGLSGFEDKRSNVELENVAKS	352
Qy	231	ERPLVY-LSXVAVGROYLYLKLETTSXSEVEAAFEALIKGVKVAQTEWKKQIDNTEVKA	289
Db	353	LSITSD	358
Qy	290	VILGSD	295
RESULT	14		
ID	030524	PRELIMINARY;	PRT; 1361 AA.
AC	030524;		
DT	01-JAN-1998 (TREMBL.REL. 05, CREATED)		
DT	01-JAN-1998 (TREMBL.REL. 05, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBL.REL. 08, LAST ANNOTATION UPDATE)		
DE	S-LAYER PROTEIN.		
OS	CRS.		
OS	CAMPYLOBACTER RECTUS.		
OC	BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;		
OC	CAMPYLOBACTER.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-314;		
RX	MEDLINE; 98187925.		
RA	WANG B.; KRAIG E.; KOLODRUBETZ D.;		
RT	"A new member of the S-layer protein family: characterization of the		
RT	crs gene from Campylobacter rectus.";		
RL	INFECT. IMMUN. 66:1521-1526(1998).		
DR	EMBL; AF010143; G2459961; -		
SO	SEQUENCE	1361 AA; 144385 MW; CAFE081F CRC32;	
Query Match	3.3%;	Score 106; DB 2; Length 1361;	
Best Local Similarity	20.4%;	Pred. No. 5.33e-01;	
Matches	38; Conservative	49; Mismatches 89; Indels 10; Gaps 10;	

 WISE RELEASE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Msearch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Aug 30 16:14:04 1999; Maspar time 18.38 Seconds
 Tabular output not generated. 544.892 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDFLAMNYDXKTL.....TISWGTTLXPQVEDKAVEND 471

Scoring table:
 Gap 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

a-geneseg35
 1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 35.164; Variance 168.905; scale 0.208

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3127	100.0	471	1	R05923	Immunogenic pneumolys	1.49e-263
2	2961	94.7	471	1	R05924	Immunogenic pneumolys	1.08e-248
3	1387	44.4	480	6	R33842	Streptolysin O varian	1.85e-108
4	1369	43.8	480	6	R33841	Soluble streptolysin	7.25e-107
5	1362	43.6	571	2	R10376	Streptolysin O deriva	3.01e-106
6	1099	35.1	371	1	R06000	Bacteriophage lambda	4.69e-83
7	112	11.8	113	3	W60952	Streptococcus pneumon	4.16e-20
8	112	11.8	113	3	R8135	Streptococcus pneumon	5.73e-20
9	108	3.5	1228	17	R77673	Listeriolysin O haemo	1.06e-01
10	108	3.5	1228	25	W2862	S-layer protein enced	1.06e-01
11	102	3.3	299	33	W60976	Bacillus steatothermo	2.63e+01
12	102	3.3	469	20	W04722	Streptococcus pneumon	2.63e+01
13	101	3.2	188	29	W55347	Automalic acyl transfe	3.05e+01
14	101	3.2	418	29	W55528	H. pylori ORF 01ce116	3.05e+01
15	100	3.2	860	8	R42962	Bovine MTP	3.54e+01
16	96	3.1	458	8	R41781	Rabbit HBI protein wh	6.33e+01

ID	Score	Query Match	Length	DB	ID	Description	Pred. No.
17	97	3.1	606	18	R33121	Tilapia prolactin rec	5.52e+01
18	97	3.1	630	18	R33120	Tilapia prolactin rec	5.52e+01
19	98	3.1	667	22	W20753	H. pylori transporter	4.76e+01
20	98	3.1	1810	16	R94563	Chicken cytoactin.	4.76e+01
21	94	3.0	114	2	R07691	Myobacterial 65KD ant	8.55e+01
22	94	3.0	355	12	R64226	Human 38kDa FK-506 bi	8.55e+01
23	94	3.0	527	33	W61278	Monamine oxidase A.	8.55e+01
24	94	3.0	527	2	R05079	Human monamine oxida	8.55e+01
25	94	3.0	625	23	W18008	Arbidopos CERI prot	8.55e+01
26	94	3.0	1057	36	W2067	HSV-2 strain SB5 Cont	8.55e+01
27	94	3.0	1114	36	W2096	HSV-2 strain SB5 Cont	8.55e+01
28	94	3.0	1114	36	W2205	HSV-2 strain SB5 Cont	8.55e+01
29	93	3.0	1786	23	W24790	P. falciparum liver s	9.88e+01
30	91	2.9	367	36	W74960	Human secreted protei	1.32e+02
31	91	2.9	480	38	W74814	Human secreted protei	1.32e+02
32	91	2.9	796	38	W85598	Cadherin-11.	1.32e+02
33	91	2.9	796	23	W25636	Human cadherin-11.	1.32e+02
34	91	2.9	796	21	W13134	Human length human cad	1.32e+02
35	90	2.9	894	8	R50007	Human MTP.	1.52e+02
36	90	2.9	964	3	R13618	C3 vegetable PEPC.	1.52e+02
37	92	2.9	3080	6	R35081	ZYMW polypeptide.	1.14e+02
38	88	2.8	84	32	W28284	Staphylococcus aureus	2.02e+02
39	88	2.8	176	22	W20335	H. pylori cytoplasmic	2.02e+02
40	89	2.8	211	1	R90412	Plasmodium falciparum	1.75e+02
41	88	2.8	531	22	W20739	H. pylori cytoplasmic	2.02e+02
42	89	2.8	649	38	W81977	Escherichia sp. B3 prot	2.02e+02
43	88	2.8	693	9	R49732	Sequence encoded by h	1.75e+02
44	89	2.8	747	39	W89585	Human ATP-binding cas	1.75e+02
45	89	2.8	757	20	W03179	Bovine poly-immunoglo	1.75e+02

ALIGNMENTS

RESULT 1
 ID R05923 standard; protein; 471 AA.
 AC R05923;
 DT 23-NOV-1990 (first entry)
 DE Immunogenic pneumolysin variant.
 OS Streptococcus pneumoniae; meningitis; bacteraemia; ds.
 KW Streptococcus pneumoniae.
 PN W09006951-A.
 PD 28-JUN-1990.
 PF 15-DEC-1989; AU0539.
 PR 15-DEC-1989; AU-001989.
 PA (PATO/) PATON J C.
 PI PATON JC, HANSMAN DJ, BOULINOIS GJ, ANDREW PW, MITCHELL TJ,
 PI WALKER JA;
 DR WPI: 90-224494/29.
 DR N-PSDB: 005270.
 PT New non-toxic, immunogenic mutants of pneumolysin - useful in
 PT protective vaccines against Streptococcus pneumoniae, and DNA
 PT sequences encoding them.
 PS Claim 5; Page 11; 25pp; English.
 CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
 CC making them useful in vaccination against pneumonia and associated
 CC meningitis, bacteraemia etc.
 SQ Sequence 471 AA:
 Query Match 100.0%; Score 3127; DB 1; Length 471;
 Best Local Similarity 94.5%; Pred. No. 1.49e-263;
 Matches 445; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Db 1 markavndflamnykklthggesianrfikenglpdfevlierkkslstntsd 60
 1 MANKAVNDFLAMNYDXKTLTHGGSINRFXKKEGNLPFXVXKRRKSLSTNTSDI 60
 OY
 Db 61 svatdsrlypgallvgeitllemptljavdrpmtysilpplassdsflgqedsn 120
 61 XVAATDSRLYPGALLVGEITLLEMPPTLJAVDRAPMTYSXKLPPLASSDSFLQVEDPSN 120
 OY
 Db 121 ssyrgavndllakwbqdyqgvnvparmqyekltahsmeglvkfgsfdektqnsldidf 180
 121 SSYRGAXXDLAKWHDYQGVNVPARMQYEXKTAHSMQGLVKFGSGDFEKGNSLDIDF 180
 OY

Db	181	nshgsekbjivtkkqiyuevsdaxkngdvtfgdtvvedlkrqgsaerplylvs	240
Qy	181	NSVHSEKKQIYVNNKQIYTVSDANKNEDVFDQTVYEDLKRGISAERPLYLSXV	240
Db	241	aygrvyvllkettksdeveaafaelikvykvapqewkqllantevxavilggdssga	300
Qy	241	AYXROYLLLETTSSXSXEVAEFALIKGVKVAPOTEWKOILDNTXVAVILGGPSSGA	300
Db	301	rvtvgtvmedlkegsrfafadpbgpisytsflndvafqgsdvyektrayn	360
Qy	301	RVVTGKVDVIEDLIQDGSFPIIDHGLPISTTISFLRDNVAFQNSIDYBTKTAA RN	360
Db	361	gdlllshsayaayqylwdelisydngkvevlprkawdngqllahftsipiknyrn	420
Qy	361	GDDLILHSGAYAAOYIITWXLSTYDHCKEVLTPKARDNGQDLTAHFETSIPLNGV RN	420
Db	421	lsvxtreclgawewrtvyehtdlprkrtisigwttllypvedkvend	471
Qy	421	LSVKIRECLGIAEMWRVRYEKTDLALVRRKTSISMGTTILPYVEDKXVEND	471

RESULT	2
ID	R05924; standard; protein; 471 AA.
AC	R05924;
DT	23-NOV-1990 (first entry)
DE	Immunogenic pneumolysin variant.
KW	Pneumolysin; vaccine; pneumonia; meningitis; bacteraemia; ds.
OS	Streptococcus pneumoniae.
EH	key
FT	misc_difference 428
FT	/label=G, A, S
FT	misc_difference 435
FT	/label=Q, D
PD	WO9006951-A.
PD	28-JUN-1990.
PF	15-DEC-1989; R00539.
PR	16-DEC-1988; AU-001989.
PA	(PATO/) PATON J C.
PI	PATON JC, HANSMAN DJ, BOUNOIS GT, ANDREW PW, MITCHELL TV,
PI	WALKER JA;
PI	WPI: 90-224494/29.
PT	New non-toxic, immunogenic mutants of pneumolysin - useful in
PT	protective vaccines against Streptococcus pneumoniae, and DNA
PT	sequences encoding them.
PS	Disclosure; P: English.
CC	Vaccines are non-toxic and antigenic to wild type pneumolysin,
CC	making them useful in vaccination against pneumonia and associated
CC	meningitis, bacteraemia etc.
SO	Sequence 471 AA.

```

Query Match 94.7% Score 2961 DB 1 Length 471:
Best Local Similarity 91.3% Pred No. 108-248;
Access 430; Conservative 5; Mismatches 36; Indels 0; Gaps 0:

Db      1 mankandfilamnydkkklthggesienrfikegnlpedefvieraikkrststnsdi 60
        |||||
        1 MANKANVNDPILAMNVDXXKLTHGSESIENRFKXGNGLPKEFVXXERKKSLSLSTSDI 60

uo      61 svlatndsrlypgaalvvadecllennpcllavdrpmtytsidpglasdsfilyedpsn 120
        |||||
        61 XYXATXDSRLYGCALIVDETYLENNPTLLADRAPMYXXKLPELASSDSFLQWEDPSN 120

Qy      121 ssvrgaxvdlkvwqbdygvnnvnapmnyekkthasmeqllvkfsgsfekngslididf 180
        |||||
        121 SSVRGAXXDLKWKHODYDQVNNVPRRXOYEXKXTHASHSEQLKVKSGSPFEKXGNSLDIDF 180

Db      181 nsyhsgeklqivnfkqilyvtvsvdavnkpgvfgdtlvevdlkqrgisaaerplyvysv 240
        |||||
        181 NSYHSGEKKIQIVNKKQIITYTSVDAKKNPGDVFQDTYVDELKRGISAERPLYISXV 240

Qy      241 aygrgyllkletsdsdevaafaalikvkvvapqewkqllidntevkavvlygddpsga 300
        |||||
        241 AYGRGYLLKLETSDSDEVAFAEALIKVKKVAPQEWKQLLIDNTEVKAVVLYGDDPSGA 300

```

```

Qy 241 AXRRQVYLKLETTSSKXMEAEPAFLIKGVKAPQJEMQKJLDNTXVAVLIGDPSGA 3000
Db 301 rvtvgkdwmedliqegsrfadnpglpjlsytsfirhvvatfnsctdyekvtayrn 3600
Qy 301 RVTVGKDWMEEDLIOEGSRFPADHPGLPISYTTSFLRDVWVATFNSDYETVLTAYRN 3600
Db 361 gdlldrgsayaagyytiffdelstfnbqgkvelrpkafdrngdclahettsiplkqavrn 4200
Qy 361 GDLLDHSSAYAAQYITWKLSTSDHOGKSEVLTPKAWDNGDVLAAHETTSIPLKQAVRN 4200
Db 421 lsvkkrregqlafgfrtvyektdipivirkttssagttlipyvedkayend 471
Qy 421 LSVKIRREGQLAMEMWRVRYEKTDILVIRKRTTISINGTTLTLPQVEDKAYEND 471

```

RESULT 3
 ID R33842 standard; Protein; 480 AA.
 AC R33842;
 DT 15-JUL-1993 (first entry)
 DE Streptolysin O variant mSLO.3/6.
 KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
 KM Streptococcus pyogenes; ASO.
 OS Streptococcus pyogenes.
 PN WO9305155-A.
 PD 18-MAR-1993.
 PE 03-AUG-1992; U06380.
 PR 30-AUG-1991; US-752428.
 PA (BECI) BECKMAN INSTR INC.
 PI Adams CW;
 WP: 93-100979/12.
 DR N-PSDB; Q38267.
 DT Streptolysin O variants produced by recombinant DNA technology -
 PT having no haemolytic activity and recognised by wild-type
 PT anti-streptolysin O antibodies, useful in diagnosis of
 PT Streptococcus pyogenes infection
 PT
 PS Claim 17; Fig 2; 80pp; English.
 CC The sequence represents a soluble variant of Streptolysin O (SLO)
 CC having no haemolytic activity. The variant SLO can specifically
 CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
 CC activity. It can be used for the determination of previous and/or
 CC current infection by Streptococcus pyogenes. It enables ASO assays
 CC to become commercially viable and avoids the risk of handling haemo-
 CC lytically active S. pyogenes wild-type SLO.
 SQ Sequence 480 AA;

Query Match	Similarity	44.4 %	Score 1387	DB 6.7	Length 480
Best Local	Similarity	40.9 %	Pred. No. 1.86e-108		
Matches 187	Conservative	100	Mismatches 169	Indels 1	Gaps 1
Db	18 indkyslnyaelewLakngetiiefyPkevgskadkfivierkkknIntpvgdisids	77			
Qy	6 VMDPLIANNYDKXKLLTHOGESIENRFKEEGNOLPXEVEXXERKKRSLSSTSDIXVAT	65			
Db	78 vldrtypaaIgaIankgfienkpdavvtrnpqkIhidIpmgmdat-vevndpIyanvt	136			
Qy	66 XDSRLYPCGALLVDSDTXLENNPITLLAVDRPMATYSXSLPGIASSDSFQVDBPSSVYRG	125			
Db	137 aIdnIvngdhdygsngtIIdpartgIetkemykskgIeaaIvnvskIdgtIgdIfkIsk	196			
Qy	126 AAXDDLAKKHQDYGGOVNVNPARXQJEXKTASHMEDQKXFGSDPEKXNSLSDIDPNVSHS	185			
Db	197 gekkymIaaykqIIfyIvsanIpmnbadYfdksvtfkelqIrkvsnaeapIIfvsnaygt	256			
Qy	186 GFKXIQIVNKKQIYTTVSVDVAKKNGDVEFDTPVEDLKORGISAEPLRYISXVAXRQ	245			
Db	257 vtrvIetIetkskndveaaIfsaalkgYcdvktngkysIilensfIavIvgqdaaehnvtk	316			
Qy	246 vYLKLETTISXSXEVAFAEALLIKGVKVAAPQTEWKQIDINTYKVAAILICGDSGGAHVTG	305			
Db	317 dIdvIvnnIkdnaatsIrkpayaIytsvIflknIagvnnIrtvevtsIeyatsI	376			
Qy	306 KYDWDIEDLIQDSRFTADHGPPISTYTSSEFLRDNVAIFPONSVDVVEIKVAIRNGDILL	365			

Query	Subject	Score	Length	DB 1	DB 2	DB 3	DB 4	DB 5	DB 6	DB 7	DB 8	DB 9	DB 10	DB 11	DB 12	DB 13	DB 14	DB 15	DB 16	DB 17	DB 18	DB 19	DB 20	DB 21	DB 22	DB 23	DB 24	DB 25	DB 26	DB 27	DB 28	DB 29	DB 30	DB 31	DB 32	DB 33	DB 34	DB 35	DB 36	DB 37	DB 38	DB 39	DB 40	DB 41	DB 42	DB 43	DB 44	DB 45	DB 46	DB 47	DB 48	DB 49	DB 50	DB 51	DB 52	DB 53	DB 54	DB 55	DB 56	DB 57	DB 58	DB 59	DB 60	DB 61	DB 62	DB 63	DB 64	DB 65	DB 66	DB 67	DB 68	DB 69	DB 70	DB 71	DB 72	DB 73	DB 74	DB 75	DB 76	DB 77	DB 78	DB 79	DB 80	DB 81	DB 82	DB 83	DB 84	DB 85	DB 86	DB 87	DB 88	DB 89	DB 90	DB 91	DB 92	DB 93	DB 94	DB 95	DB 96	DB 97	DB 98	DB 99	DB 100	DB 101	DB 102	DB 103	DB 104	DB 105	DB 106	DB 107	DB 108	DB 109	DB 110	DB 111	DB 112	DB 113	DB 114	DB 115	DB 116	DB 117	DB 118	DB 119	DB 120	DB 121	DB 122	DB 123	DB 124	DB 125	DB 126	DB 127	DB 128	DB 129	DB 130	DB 131	DB 132	DB 133	DB 134	DB 135	DB 136	DB 137	DB 138	DB 139	DB 140	DB 141	DB 142	DB 143	DB 144	DB 145	DB 146	DB 147	DB 148	DB 149	DB 150	DB 151	DB 152	DB 153	DB 154	DB 155	DB 156	DB 157	DB 158	DB 159	DB 160	DB 161	DB 162	DB 163	DB 164	DB 165	DB 166	DB 167	DB 168	DB 169	DB 170	DB 171	DB 172	DB 173	DB 174	DB 175	DB 176	DB 177	DB 178	DB 179	DB 180	DB 181	DB 182	DB 183	DB 184	DB 185	DB 186	DB 187	DB 188	DB 189	DB 190	DB 191	DB 192	DB 193	DB 194	DB 195	DB 196	DB 197	DB 198	DB 199	DB 200	DB 201	DB 202	DB 203	DB 204	DB 205	DB 206	DB 207	DB 208	DB 209	DB 210	DB 211	DB 212	DB 213	DB 214	DB 215	DB 216	DB 217	DB 218	DB 219	DB 220	DB 221	DB 222	DB 223	DB 224	DB 225	DB 226	DB 227	DB 228	DB 229	DB 230	DB 231	DB 232	DB 233	DB 234	DB 235	DB 236	DB 237	DB 238	DB 239	DB 240	DB 241	DB 242	DB 243	DB 244	DB 245	DB 246	DB 247	DB 248	DB 249	DB 250	DB 251	DB 252	DB 253	DB 254	DB 255	DB 256	DB 257	DB 258	DB 259	DB 260	DB 261	DB 262	DB 263	DB 264	DB 265	DB 266	DB 267	DB 268	DB 269	DB 270	DB 271	DB 272	DB 273	DB 274	DB 275	DB 276	DB 277	DB 278	DB 279	DB 280	DB 281	DB 282	DB 283	DB 284	DB 285	DB 286	DB 287	DB 288	DB 289	DB 290	DB 291	DB 292	DB 293	DB 294	DB 295	DB 296	DB 297	DB 298	DB 299	DB 300	DB 301	DB 302	DB 303	DB 304	DB 305	DB 306	DB 307	DB 308	DB 309	DB 310	DB 311	DB 312	DB 313	DB 314	DB 315	DB 316	DB 317	DB 318	DB 319	DB 320	DB 321	DB 322	DB 323	DB 324	DB 325	DB 326	DB 327	DB 328	DB 329	DB 330	DB 331	DB 332	DB 333	DB 334	DB 335	DB 336	DB 337	DB 338	DB 339	DB 340	DB 341	DB 342	DB 343	DB 344	DB 345	DB 346	DB 347	DB 348	DB 349	DB 350	DB 351	DB 352	DB 353	DB 354	DB 355	DB 356	DB 357	DB 358	DB 359	DB 360	DB 361	DB 362	DB 363	DB 364	DB 365	DB 366	DB 367	DB 368	DB 369	DB 370	DB 371	DB 372	DB 373	DB 374	DB 375	DB 376	DB 377	DB 378	DB 37
-------	---------	-------	--------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	-------

	CC	treatment or prevention of infection, specifically caused by S.
	CC	pneumoniae (particularly meningitis) but possibly also Helicobacter
	CC	Pylori (ulcers and gastric cancer). It may be of particular
	CC	use before insertion of an in-dwelling device or any other
	CC	Invasive procedure. The protein, or nucleic acid encoding
	CC	it, can also be used in vaccines to induce a cellular
	CC	and/or humoral immune response, or to screen for other
	CC	antibacterials. The DNA may also contain flanking sequences
	CC	that are potential sources of control elements for bacterial
	CC	gene expression. Detecting a sequence encoding the protein
	CC	can be used diagnostically, e.g. to detect a mutation for
	SO	Serotyping or classifying infectious agents. Sequence 113 AA;
Dd	Query Match	11.8%; Score 370; DB 33; Length 113; Best Local Similarity 43.2%; Pred. No. 4,16-20;
OY	Matches 48; Conservative 31; Mismatches 31; Indels 1; Gaps 1;	
Dd	2 kdkllilqhsagyaiaarsltweeypyvdkgnyvrvshswegnrgtafvlnlpikennm 61 :: OY RNSGDLIDSGAGVAYQXYITWXLSTYDHQGKEVLFRADMRNDGLTAFTTSPILKNV 418	
Oy	62 nrlvykktqtgllwnrtgltyenrrplaaqbkhkihwftlnskrsddv 112 OY 419 RNLSKIRECTGIAMEMWRIVTEKTDLAVRKRTISIMWTIIPDY-EDKV 468	
RESULT	8	
ID	R88135 standard; peptide: 11 AA.	
AC	R88135:	
DT	29-AUG-1996 (first entry)	
DE	Listeriolysin O haemolytic domain streptolysin O pneumolysin;	
KM	Listeriolysin O haemolytic domain streptolysin O pneumolysin;	
KW	soluble molecular complex specific targeted polynucleotide;	
KX	endosome lysins bacterial component cytoplasm therapeutic;	
KV	in vivo ex vivo in vitro delivery cell specific hemolytic.	
OS	Synthetic.	
PN	WO960792-A1.	
PD	11-JAN-1996.	
PF	27-JUN-1995; U08091.	
PR	29-JUN-1994; US-567710.	
PS	07-JUN-1995; US-484009.	
PA	(TARGET-) TARGETECH INC. (UYCO-) UNIV CONNECTICUT.	
PT	Carnichael E Spitalny GL Wu CH Wu GV Zhang Y;	
DR	WPt: 96-077502/08.	
PI	Soluble molecular complex for specific targeting of PN to cell -	
PT	internalises polynucleotide (PN) into endosome, then lyses endosome	
PS	to release PN into cell's cytoplasm	
Example 3:	Page 13: 23pp; English.	
CC	The present peptide is a listeriolsyin O (LLO) haemolytic domain	
CC	(HD) fragment which is 100% conserved in the LLO homologues	
CC	streptolysin O, and pneumolysin A. A 201 bp fragment encoding the LLO	
CC	HD can be used instead of the full length LLO coding sequence in a	
CC	soluble molecular complex, for the specific targeting of a	
CC	polynucleotide (PN) to a cell. The complex comprises a PN, a	
CC	carrier comprising a PN binding agent and a cell specific binding	
CC	agent, which binds to a surface mol. of the cell and its	
CC	internalised into an endosome, and a bacterial component i.e. the	
CC	LLO HD fragment), which lyses the endosome to release the	
CC	PN into cell's cytoplasm. The complex can be used therapeutically	
CC	for enhanced in vivo, ex vivo or in vitro delivery of a PN to a	
CC	specific cell.	
SO	Sequence 11 AA;	
Query Match	3.6%; Score 112; DB 17; Length 11;	
Best Local Similarity	100.0%; Pred. No. 5,73e+00;	
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Dd	1 ectqlawewr 11 	
OY	427 ECTGLAMEWMR 437	

PT negative bacterium - or new sbs-B gene in any host, also new
PT recombinant proteins containing heterologous inserts, e.g.
PS epitope(s) useful as vaccines and adjuvants
PS Claim 1, Pages 9-14; 31pp; German.
CC The present sequence is the Bacillus stearothermophilus PV72
CC S-layer protein, sbs-A. S-layer structures can be used as vaccines
CC or adjuvants, particularly when they include a bacterial ghost that
CC may contain additional epitopes in its membrane. Other uses of
CC recombinant sbs-A, depending on the nature of the inserted peptide,
CC are as an universal carrier for biohybridated reactants for use in
CC immunological or hybridisation assays (the insert is streptavidin),
CC to induce immune responses (epitopes), as a reagent for removing
CC cytokine or toxin from serum (antigenic epitopes), as a molecular
CC spinning nozzle (polyhydroxybutyrate synthase) and as a molecular
CC laser (luciferase).
SQ Sequence 1228 AA;

Query Match 3.5%; Score 108; DB 25; Length 1228;
Best Local Similarity 23.5%; Pred. No. 1,066+01;
Matches 39; Conservative 43; Mismatches 74; Indels 10; Gaps 10

Dd 200 ydltv-amkar-ev-qdavkagnldkaavdqnglypxvtadfkelt-evakkalda 255
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:
Qy 200 YIVSVDAKPNPDVVDQDVYTEDLDRGSIASERPLVYSXVAXYRQYLKLETTISXSEV 259
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:

Dd 256 deaaltpkvesys-aifnqkaveltaypvgnt-lklqlsaanedlvnnvtviyrxvdg 313
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:
Qy 260 -EAAFEALIKGVKAPQIEWKQI-LDNTXKAVAILGGPSSGAAVVVGKDMDVEDLIQEG 317
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:

Dd 314 nlpfalfatdvsistcdgttlivdstpfenttekykvvvkgikdxng 359
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:
Qy 318 S-RFTADHPGLPISTYTSFLRDNVAFQNSTDY-VETKVATVANG 361
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:

RESULT 11
ID W60976 standard; Protein; 299 AA.
AC W60976;
DT 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae encoded polypeptide.
KW coding region; ORF: open reading frame; antibacterial;
KM Infection: prevention; meningitis.
OS Streptococcus pneumoniae.
PN W09819689-Al.
PD 14-MAY-1998.
PE 27-OCT-1997; U19226.
PR 01-NOV-1996; US-029930.
PA (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
PI Black MT, Hoods JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PN;
DR WPI: 98-286586/25.
DR N-PSDB: V37371.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
PT Of meningitis
PT Claim 11, Page 71-72; 130pp; English.
PS The sequence is that of the polypeptide encoded by a region isolated
PS from S. pneumoniae which shows homology to nicotinate-nucleotide pyro-
PS phosphorylase. It, or agonists of it, may be useful as an anti-
PS bacterial for treatment or prevention of infection, specifically caused
PS by S.pneumoniae (particularly meningitis) but possibly also Helicobacter
PS Pylori (ulcers and gastric cancer). It may be of particular
PS use before insertion of an in-dwelling device or any other
PS invasive procedure. The protein, or nucleic acid encoding
PS it, can also be used in vaccines to induce a cellular
PS and/or humoral immune response, or to screen for other
PS antibacterials. The DNA may also contain flanking sequences
PS that are potential sources of control elements for bacterial
PS gene expression. Detecting a sequence encoding the protein
PS can be used diagnostically, e.g. to detect a mutation for
PS serotyping or classifying infectious agents.
SQ Sequence 299 AA;

```

Query Match      3.3%; Score 102; DB 33; Length 299;
Best Local Similarity 20.4%; Pred. No. 2.63e+01;
Matches          21; Conservative         33; Mismatches    43; Indels     6; Gaps       6;

Dc   38 dycnaiafhgqakvslfakagvaig-ltvfgvfllfaevtfqphgtkdgrlts 96
        |::|::||::||::||::||::||::||::||::||::||::||::||::||:
Cy   14 NYDXKKLLTHQGEIEENRFKKEHQJLPEFVXXEKRRKSRLSTNTS-DIXXATXSRLPY 72
        ||::||::||::||::||::||::||::||::||::||::||::||::||::

Db   97 gdl-vl-el-igsrvsliltoervalnlnqlpslaamtayve 136
        ||::||::||::||::||::||::||::||::||::||::||::||::||::
Cz   73 GALLVDETFKLNNPPTLLAVDRAPMTSXSLPGLAS-SDSFLQ 114

RESULT 12
ID W04722 standard; Protein; 469 AA.
AC W04722;
DT 06-FEB-1997 (first entry)
DE Aromatic acyl transferase of Gentiana triflora var. japonica.
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KM plants; acylation; colour; tone; colouration; colour change;
KV Gentiana triflora; Petunia hybrid; Perilla ocimoides;
SC Senecio cruentus; Lavandula angustifolia.
OS Gentiana triflora var. japonica (Clone pCARF4).
PN W09625300-AI.
PF 22-AUG-1996.
PR 16-FEB-1996; J00348.
PR 17-FEB-1995; JP-067159.
PR 29-JUN-1995; JP-196915.
PR 30-JAN-1986; JP-046534.
(SUNR ) SUNTORY LTD.
PA Ashikari T., Fujiwara H., Fukui Y., Kusumi T., Mizutani M.;
PI Nakao M., Tanaka K., Yonekura K.;
PI WP1_96:393401/39.
DR N-PDBB. I37308.
DI DNA coding for aromatic acyl transferase - for transforming plants
PI which produce anthocyanin pigments and thus altering colour tone,
PT e.g. of flowers
PS Claim 4; Page 33-57; 94pp; Japanese.
CC Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in I37308-I37313.
SQ Sequence 469 AA;

Query Match      3.3%; Score 102; DB 20; Length 469;
Best Local Similarity 19.8%; Pred. No. 2.63e+01;
Matches          32; Conservative         49; Mismatches    70; Indels    11; Gaps     11;

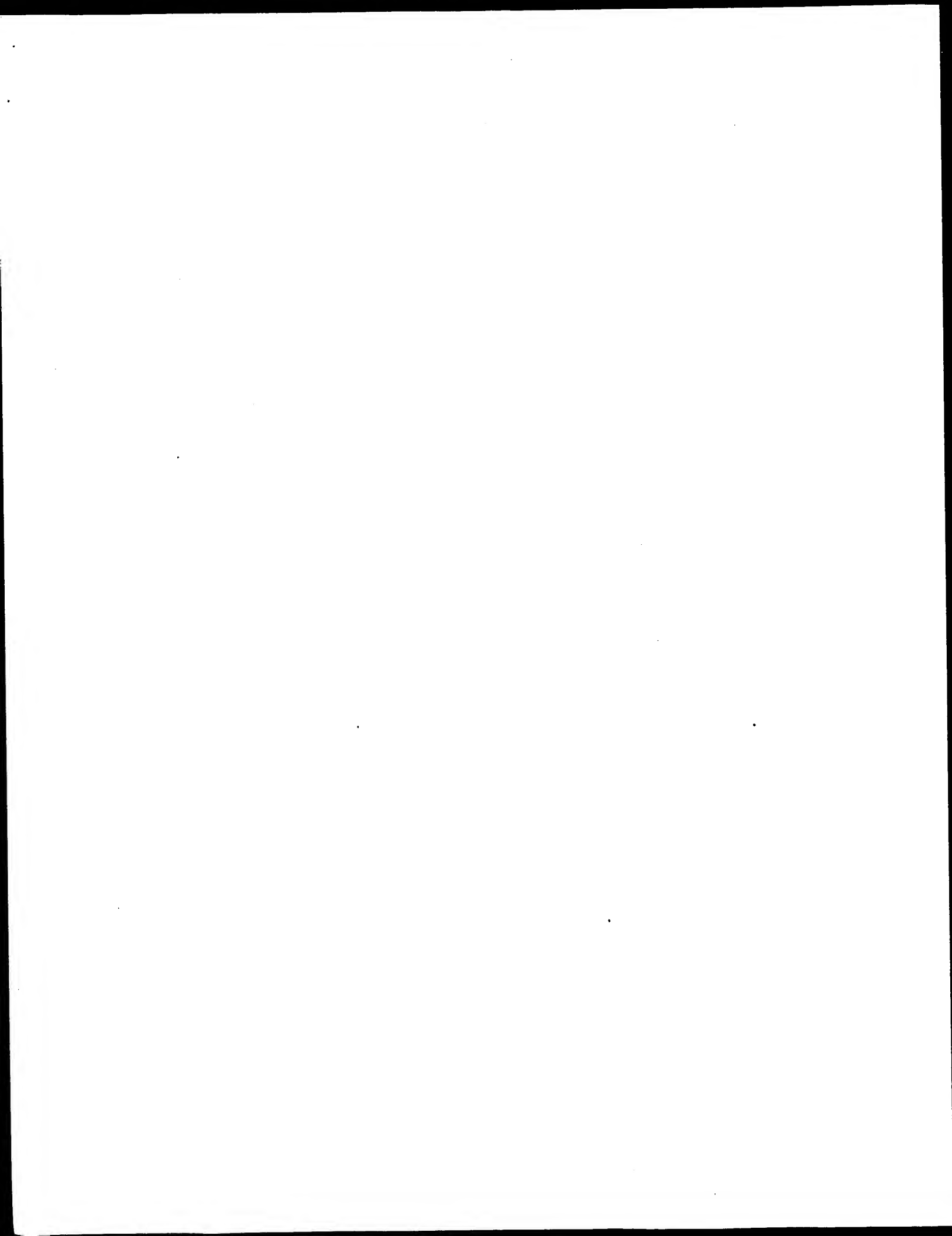
Dc   94 pkfyrsrdgdstiltivaesddfdylyghnl-vsndhlglfymrpvitmqdy-kv 150
        ::||::||::||::||::||::||::||::||::||::||::||::||::
Cy   146 ARXYAEKYTAHSHMEQLKFVGSDFE-KXGNSLDIDFNVSHEGXKI-QIVNKKQYTVTS 203
        :||::||::||::||::||::||::||::||::||::||::||::||::
Db   151 iplvavqvfvonrgjaavlatahsjadsksfvmfinawayink-fgr-dadllaanlip 208
        ||::||::||::||::||::||::||::||::||::||::||::||::||:
Cz   204 VDAAVKMPGDVPFDITVEIDL-QRGISARPLV-YISXAAYRYLYLKLETTSSXEVEA 261
        ||::||::||::||::||::||::||::||::||::||::||::||::||:
Db   209 sfirsiikdylgleetfwne-mqdvlmfmrfsrkypprfnkxv 249
        |:|::||::||::||::||::||::||::||::||::||::||::||::
Cz   262 AFE-ALIKGVKAPOPTEMKOILDINTFYKVAVLIGDPSPSGGARV 302
        ||::||::||::||::||::||::||::||::||::||::||::||::||:

RESULT 13
ID W55347 standard; Protein; 188 AA.
AC W55347;
DT 17-JUN-1998 (first entry)
DE Cytosolic ORF O1cel116lorfil protein.
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KM identification; binding compound; bacteria; life cycle; activator;
KM inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS Helicobacter pylori.
```

```

PN MO9737044-AL.
PD 09-OCT-1997.
PD 27-MAR-1997; U052223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR ) ASTRA AB.
PI Alm RA, Smith D;.
DR WPI: 97-503122/46.
DR N-PSDB: V24756.
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
infection and for diagnosis of H. pylori infection
PS Claim 14, Page 571; 1145pp; English.
CC This sequence is a H. pylori protein of unspecified function.
CC The protein may be used in a vaccine to prevent or treat H. pylori
infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of H. pylori mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of H. pylori-specific antigens. The genomic sequence of
CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely H. pylori antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or exported
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from H. pylori by PCR
CC amplification for recombinant polypeptide production, e.g. in E. coli
CC hosts.
SQ Sequence 188 AA;
Query Match 3.2%; Score 101; DB 29; Length 188;
Best Local Similarity 32.8%; Pred. No. 3,05e+01;
Matches 20; Conservative 20; Mismatches 18; Indels 3; Gaps 3;
Db 95 psmlrdsvsledskkr-lnaikdlifhnkaf-rqlqklntplkal-vesagkgdefna 151
: :|::||| | :|: ||: ||| | : ||| ::|:
QY 210 PEDVQRDVIYVEDLKRGSGISERPLVYSXVAAXRQYLLETFSTXSXEVEAFALLING 269
Db 152 1 152
QY 270 Y 270
RESULT 14
ID W55528 standard; Protein; 418 AA.
AC W55528;
DI 02-JUL-1998 (first entry)
DE H. pylori ORF 29gc0321_24336712.fl5 cellular protein.
KM Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS Helicobacter pylori.
ON WO9737044-AL.
PD 09-OCT-1997.
PD 27-MAR-1997; U052223.
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 02-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR ) ASTRA AB.
PI Alm RA, Smith D;.
DR WPI: 97-503122/46.
DR N-PSDB: V24937.
PT Helicobacter pylori nucleic acid sequences and encoded

```

 Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 (1M)

 Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSEARCH protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Aug 30 16:12:54 1999; MasPar time 20.74 Seconds
 Tabular output not generated. 909,898 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDFILAMNDXXKL.....TISMGITLYPQVEDKEND 471

Scoring table: PAM 150
 Gap 11
 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 50.051; Variance 108.607; scale 0.461

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	3127	100.0	471	2	A28568	0.00e+00
2	1556	49.8	500	2	B43577	3.55e-265
3	1475	47.2	529	2	A43505	1.78e-249
4	1468	46.9	529	2	S24231	4.04e-248
5	1460	46.7	528	2	S22341	1.43e-246
6	1416	45.3	501	2	A37858	4.66e-238
7	1414	45.2	530	2	S22340	1.14e-237
8	1379	44.1	571	2	A43507	6.61e-231
9	1298	41.5	485	2	I39863	2.84e-215
10	1298	41.5	485	2	I39863	2.84e-215
11	1209	36.7	96	2	S47298	3.51e-58
12	113	3.6	265	2	S60947	1.01e-14
13	112	3.6	527	2	S03974	2.77e-01
14	112	3.6	744	2	A70385	3.63e-01
15	114	3.6	1361	3	T03415	3.63e-01
16	109	3.5	218	2	F64461	2.11e-01
17	108	3.5	462	1	S00552	8.09e-01
18	108	3.5	1228	2	I40468	1.05e+00
19	107	3.4	364	2	J01188	1.05e+00
20	106	3.4	426	2	A70444	1.37e+00
21	105	3.4	776	1	RGECAK	1.78e+00
22	105	3.4	780	1	S77652	2.30e+00
23	102	3.3	365	2	F71665	4.95e+00

24	103	3.3	446	2	E69899	L-amino acid oxidase	3.84e+00
25	102	3.3	514	1	DWE87T	threonine dehydratase	4.95e+00
26	103	3.3	722	2	S64492	hypothetical protein	3.84e+00
27	102	3.3	878	2	F64425	valine--tRNA ligase	4.95e+00
28	99	3.2	286	2	S55901	probable tagatose 1,6	1.04e+01
29	100	3.2	287	2	C71207	hypothetical protein	8.16e+00
30	100	3.2	349	2	H64456	hypothetical protein	8.16e+00
31	100	3.2	362	2	B70876	probable aminotransfe	8.16e+00
32	99	3.2	366	2	E69375	ABC transporter, ATP-	1.04e+01
33	101	3.2	418	2	G71952	hypothetical protein	6.36e+00
34	100	3.2	427	2	S75210	glycine hydroxymethyl	8.16e+00
35	101	3.2	501	2	A69415	conserved hypothetical	6.36e+00
36	100	3.2	502	2	UQ2341	arginine decarboxylas	8.16e+00
37	100	3.2	511	2	S58322	nucleolar protein NOP	8.16e+00
38	99	3.2	877	2	S58824	probable membrane pro	1.04e+01
39	100	3.2	887	2	A46764	microsomal triglyceri	8.16e+00
40	100	3.2	899	2	A44879	retinoblastoma protei	8.16e+00
41	101	3.2	1026	2	A49750	beta-lactamase (EC 3.	6.36e+00
42	99	3.2	1035	2	A64686	cation efflux system	1.04e+01
43	100	3.2	1123	1	WMBEH7	UL37 protein - human	8.16e+00
44	99	3.2	3027	2	J01917	polyprotein - parvini	1.04e+01
45	100	3.2	5762	2	A41819	proline-rich peptides	8.16e+00

ALIGNMENTS

RESULT 1
 ENTRY A28568 #type complete
 TITLE pneumolysin - Streptococcus pneumoniae
 ORGANISM #formal name Streptococcus pneumoniae
 DATE 19-Nov-1988 #sequence_rev19-Nov-1988 #extl_change

ACCESSIONS
 #authors A28568
 #journal Walker, J.A.; Allen, R.L.; Falmagne, P.; Johnson, M.K.;
 Infect. Immun. (1987) 55:1184-1189

#title Molecular cloning, characterization, and complete nucleotide
 sequence of the gene for pneumolysin, the
 sulphydryl-activated toxin of Streptococcus pneumoniae.

#cross-references M01D:87193109
 #accession A28568

#molecule_type DNA
 #residues 1-471 #label WAL

REFERENCE
 #authors Mitchell, T.J.; Mendez, F.; Paton, J.C.; Andrew, P.W.;
 Boulnols, G.J.

#journal Nucleic Acids Res. (1990) 18:4010
 #title Comparison of pneumolysin genes and proteins from
 Streptococcus pneumoniae types 1 and 2.

#cross-references EMBL:X52474; NID:947403; PID:947404
 #superfamily dipeptide transport protein

FEATURE
 #residues 1-471 #label MWT

CLASSIFICATION
 #product pneumolysin #status predicted #label MAT

SUMMARY
 #length 471 #molecular_weight 52899 #checksum 8274

Query Match 100.0%; Score 3127; DB 2; Length 471;
 Best local similarity 94.5%; Pred. No. 0.00e+00;
 Matches 445; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

DB 1 MANKAVNDFILAMNDXXKLTHOGESIFENRFEKGNOLPEEFVIRKRSTSTNSDI 60
 QY 1 MANKAVNDFILAMNDXXKLTHOGESIFENRFEKGNOLPEEFVIRKRSTSTNSDI 60
 DB 61 SYTANDSRLPGALVYDFELLNNPTLLAVDRAPMYTSIDLPGLASSDSFLOVEDPSN 120
 QY 61 YXATXDSRLPGALVYDFELLNNPTLLAVDRAPMYTSIDLPGLASSDSFLOVEDPSN 120

ENTRY	2
TITLE	B43577 #type complete
ORGANISM	perfringens Clostridium perfringens
DATE	#format_name Clostridium perfringens
	03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change
	13-Sep-1998
ACCESSIONS	B43577; A34951; A60922
REFERENCE	A43577
#authors	Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
	Infect. Immun. (1991) 59:137-142

```

#title      An upstream regulatory sequence stimulates expression of the
            perfringolysin O gene of Clostridium perfringens.
#cross-references  MIMD:91099951
#accession  B43577
#molecule_1-type  DNA
#residues    1-500 #label SHI
#cross-references  GB:M81080; NID:g144885; PID:g144886
#note        translation of the nucleotide sequence is not complete
REFERENCE
#authors    Tweten, R.K.
#journal    Infect. Immun. (1988) 56:3235-3240
#title      Nucleotide sequence of the gene for perfringolysin O
            of Clostridium perfringens strain 1302A.

```

```

#cross-references (theta-toxin) from Clostridium perfringens: significant
#cross-references MUID:89032623 homology with the genes for streptolysin O and pneumolysin
#accession A34951
#molecule_type DNA
#residues 1-125, 'EA', 129-500 ##label TWE
#cross-references GB:M36704; NID:G144883; PID:G144884
REFERENCE
#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3228-3234
#title Cloning and expression in Escherichia coli of the
#title perfringolysin O (theta-toxin) gene from Clostridium
#title perfringens and characterization of the gene product.
#cross-references MUID:89032622
#accession A60922
#molecule_type protein
#residues 29-45 ##label TW2
#experimental_source ATCC 13124

```

	GENETICS	
gene	pfo; pfoA	
FEATURE		
1-78	#domain signal sequence	#status predicted #label SIG
29-500	#product pefringolysin O	#status experimental #label
	NAT	
SUMMARY	#length 500	#molecular-weight 55799 #checksum 4821

```

Db 35 N0SIDGJSSLSYENEVLASNGKISFPFKGKGAQKFTIVEORKSFLTSPVDISI 94
Qy 3 NKAYNDFILAMNDXXKLLTHQGESIENRXKJEEBOLPXEYVAXXKKKSSLTNSDIXV 62
Db 95 IDAVNDRTPGALOTLADKAFENRPTILMWKRKININIDLPGLKGENS- IKVDPTYGK 153
Qy 63 XAIXDRLKPGALLVVEDELXLENPPTLLAVDRAPMYXSXLPGLASDSDFLVEDPENSNS 122
Db 154 VSGAIDELYSKNNKXSSHTLPRATQYSEBMYSKSOISALTANVAXYLENSLGDVNA 213
Qy 123 VRGAXXDLLAKHMODGVGVNNPAPXQJEXKTAHSMQOLVYFGSGFPEKXGNSLDIDNS 182
Db 214 VANNKRWMLIAWQJFYFYNSADLPKNSDLEDDSVTFNDLKQGVSNAPPLMNVAY 273
Qy 183 VHSEKKIQIVNNKQIYTVSVDAKNNGVDFODTVHEDLKRGISARPLVYISXVAY 242
Db 274 GRITYAKLETTSSSKVQAFAALLKNDIKNSOQYDIENSSFTAVYILGDPAGEHNKX 333
Qy 243 XROYVLELTTSXSKVEAEFALLIKGVKVPJOTEMQOILDNKXKRAVILGDPSSGAV 302
Db 334 VTQDFEIRKIVINDNFTSTKPNAPRISYTSVFTKDNSVAVAHNKTDYLETTSTEXSKG 393
Qy 303 VTGKVDWVEDLIEGSRFADHPGLPISTTSFLRDNVAVAFQSDLSVETTKVATYRND 362
Db 394 INDDHSAVYAQGEVAVMDVSYDKGNGVNLHKHMDGVQOYKTHAYTVLEPANRNR 453
Qy 363 LLDHSSAVYAQYITWXLSTHOGKEVLELPKANDRNGQDLTHFTTSLPLGNVRNLS 422
Db 454 IKARECTGLAMWMDVISEYDPLTNININSIWGTTLLP 493
Qy 423 VKIRECTGLAMWMDRYEKRTDLXVRKRTTISIWGTTLLP 462

```

Accessions	Date	Organism	Title	Entry	Result
A43505	17-Mar-1999	<i>Listeria monocytogenes</i>	21-Oct-1992 #sequence_revision	A43505	3
A47606	21-Oct-1992	<i>Listeria monocytogenes</i>	21-Oct-1992 #text_change	A47606	
A61079	21-Oct-1992	<i>Listeria monocytogenes</i>	21-Oct-1992 #sequence_revision	A61079	
A505306	17-Mar-1999	<i>Listeria monocytogenes</i>	21-Oct-1992 #sequence_revision	A505306	

#cross-references MUID:8936684
 #journal #title
 #journal #title
 #cross-references MUID:88153053
 #accession A43505
 #status preliminary
 #molecule-type DNA
 #residues 1-529 #label MEN
 #cross-references GB:M2119; NID:g149652; PID:g149653
 #note this sequence is derived from a strongly hemolytic strain, serotype 1/2c
 REFERENCE
 #authors Domann, E.; Chakraborty, T.
 #journal Nucleic Acids Res. (1988) 17:6406
 #title Nucleotide sequence of the listeriolysin gene from a *Listeria monocytogenes* serotype 1/2a strain.

```
#accession      S05306
#molecule_type DNA
#residues       1-529 #label DOM
##cross-references EMBL:X1517; NID:g44106; PID:g44107
##experimental_source strain EGD
#note           this sequence is derived from a weakly hemolytic strain,
                  serotype 1/2a

REFERENCE
AA7606
```


[illegible]

ORGANISM	#formal_name	Listeria monocytogenes
#variety	strain 12067	
DATE	22-Nov-1993	#sequence_revision 10-Nov-1995 #text_change 13-Sep-1998
ACCESSIONS	S24231	
REFERENCE	S24231	
#authors	Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.	
#journal	Infect. Immun. (1991) 59:3945-3951	
#title	Listeria monocytogenes isolates can be classified into two major types according to the sequence of the listeriolysin gene.	
#cross-references	MUID:92040062	
#accession	S24231	
#status	nucleic acid sequence not shown; translation not shown	
#molecule_type	DNA	
#residues	1-529	#label_RAS
#cross-references	EMBL:X60035; NID:g44110; PID:g44112	
#experimental_source	strain 12067, serotype 4b	
#note	the nucleotide sequence was submitted to the EMBL Data Library, June 1991	
GENETICS	listA	
CLASSIFICATION	superfamily dipeptide transport protein	
KEYWORDS	virulence factor	
FEATURE		
1-25		
26-529		
SUMMARY		
Query Match	46.9%; Score 1468; DB 2; Length 529;	
Best Local Similarity	40.6%; Pred. No. 4,04e-248;	
Matches 189; Conservative 115; Mismatches 161; Indels 1; Caps 1;		
Db	61 IDKYIOGLDLYKKNVLYHGDVAVNPPRGYKDGNGEYIVAEKKRSINONNDIQVYNA	120
Qy	6 VNDFIANNYYXXLLTHQGESINRRKXKBNQLPXEFVXXEKKRKLSTNTSDIYXAT	65
Db	121 ISSLTYPALYKANSELVENOPDLYPKRDSLTSLIDLPGMTNODNKIYKNAKTSVNN	180
Qy	66 XDSRLYPALYVETXLENNPRLTLLADRAPMYXXLPGLASDSPLOYEDPSNSVRC	125
Db	181 AVNTLVEWNRKYAOAVPNYSAKIDYDDEMAYSQJIAFGIAFRKVNNSLVANGAIS	240
Qy	126 AXHDLAKHODIQOVN-NVPAKQYQKXTHASHSEQKLVKFGDFEKKXGSLDIDRFSV	184
Db	241 EGMQEEVYSKQIYNNVNEPTRPSPRFGKAVTKQELQALQVNAENPPAYISVAAGR	300
Qy	185 SGEKKIQIVNKKQIYTVSDAVANPGDVFQDPTVEDLQNRGISMERPLVYISYAAIKR	244
Db	301 QVYTLKLSNHSITVKAFDAVSGKVSQDVELTNIKKSKRAVIYGGSADEVQIID	360
Qy	245 QVYTLKLETTSSXVEAFELKIGVAVAPQOTEMKQILDTYKAVAILGGDPSSGARVY	304
Db	361 GNLGDLRLDKGATFENRETPGVFIATYTFMLKDNELAVKNNSEVITETSKAYTGKIN	420
Qy	305 GKVDMEEDLQOEGSRFADHGLPISTTISFLDNVATQNSDTEVERKYVIRMGDL	364
Db	421 IDHSGGYAOGNISWDEINTDPGENELVQHKNNSENKSKLAHFTSSIYLPGNARNINY	480
Qy	365 LDHSGAIVAAOYIITXWELSYDHQGEKLETPKARDRNGODLTAFITTSIPKGNVRLSVK	424
Db	481 AKETGTAMEMWRTVIDRNLPLKNNNISIWGTTLYPKXNSVDN	526
Qy	425 IRETGTAMEMWRTYKEDLDLXKRTKTSIWGTTIIPQVEDKVEN	470
RESULT	5	
ENTRY	S22341	#type complete
TITLE	ivanolysin precursor - Listeria ivanovii	
ORGANISM	07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Aug-1998	

```

ACCESSIONS      S22341; S36683
REFERENCE        S22340
#authors        Haas, A.; Dumbsky, M.; Kretz, J.
#journal        Blochm., Biophys. Acta (1992) 1130:81-84
#title          Listeriolysin genes: complete sequence of 110 from Listeria
                  ivanovii and of 180 from Listeria seeligeri.
#cross-references MUID:92182018
#accession      S22341
                  ##molecule_type DNA
                  #residues 1-528 ##label HAS
                  ##cross-references EMBL:X60461
                  #note the authors translated the codon ACA for residue 331 as
                           Val
REFERENCE        S36683
#authors        Kretz, J.
#submission      submitted to the EMBL Data Library, July 1991
#accession       S36683
                  ##molecule_type DNA
                  #residues 1-319,'T',321-528 ##label KRE
                  ##cross-references EMBL:X60461
GENETICS
#gene            110
CLASSIFICATION . #superfamily dipeptide transport protein
FEATURE
1-24             #domain signal sequence #status predicted #label SIG\
25-528           #product ivanovylisin #status predicted #label MAT
SUMMARY          #length 528 #molecular-weight 58511 #checksum 6874

Query Match      46.7%: Score 1460; DB 2; length 528:
Best Local Similarity 40.8%; Pred. No. 1,438-246; Indels 1; Gaps 1
Matches 190; Conservative 116; Mismatched 159;

Db 60 IDQYIOGLDYDKNNILVYDGEAVKVPKAGYKEGNOYIYVEKKRSINQNNADIYVNS 119
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 6 VNDFTAMNYDXKKLLTHQGESIERFKYEGNOQLPEFVXXEKKRRSRSTNSDIYXAT 65
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 120 IASLTPGALPKANSELVENQPDVLPVKRDSVTLSIDLPGWNHNELVQGNATKRSIN 179
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 66 XDSRLYPGALLVVDFTXLENNPTLLAVRAPATYSXSLPGLAASSDFLQVEDPSNVSYRG 125
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 180 GVNLTIVDRMNNKYSSELYPNISAKIDYDDEMAVSESOQLAKFGAFAKAVNNSLNWFAGIS 239
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 126 AXDXLLAKRHODYGOVN-NVPRARXOYEKKXTHASMQLKVKFGSDPEKKGNSLIDFNFSVH 184
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Oy 240 ECKVOEEVYNFKQIITYTVANNEPTSPSRFFGKSVTKENLDLGVNAENPRAIYSSVACGR 299
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 185 SGEKKIIOYVNNKQIITYTVSDAVKNGVGFDTYVEDLKGKGISAEPRPLYISXVAAYXR 244
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 300 DLFVKLSTSHSTRKAFDAFAKFSKVGDPDELNIITONASFRAVIYSGSKADEVEIID 359
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 245 GYVILKLETTSKXKEVBAEAFALIKGVKAPQPEEMKQIIDNTYKRAVILIGDPSSGARVYT 304
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Oy 360 GDLKKRILDKOGANPEKKNPGVPIAYTTNPLKDNOLAVNNKSEIETTSKANSYDGKIN 419
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 305 GAYDWMEDDIQGSRFTADHPGLPISTYSFELRDVVVATFQNSTDYVEFKTAYRNQDGL 364
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 420 LDHSGAYVAREVTVDEYVDYDANGNNVYHKKXSENDOKLAHETTSYIPLGNANINIH 479
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 365 LDHSGAYVAYQYITTXKELSYDHQGEVELPKAMDRNGODLTAHFTTSLPLKGNVNNLSVK 424
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db 480 AKECTGLAEWNRVTVDDNRDLPLVKRNVCINGTLLYPAYSDTVDN 525
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy 425 IRECTGLAEWNRVTEYKTDKLIVKRRITISMGTTLYPQVEDKVEN 470
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT          6
ENTRY           A37858 #type complete
TITLE           alveolysin - Bacillus alvei
ORGANISM        #formal_name Bacillus alvei
DATE            28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change
                  09-Sep-1997
ACCESSIONS      A37858
REFERENCE        A37858

```

[illegible]


```

ALTERNATE_NAMES  hypothetical protein 05050; hypothetical protein Y0R50-10
ORGANISM          #format name Saccharomyces cerevisiae
DATE              15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change
                  05-Jun-1998

ACCESSIONS
REFERENCE
#authors          G60947; S67113; S71722
#submission       G60938
#description       Submitted to the EMBL Data Library, October 1995
#accession        Sequence and analysis of a 33 kb fragment from the right arm
#molecule_type   of chromosome XV of the yeast Saccharomyces cerevisiae.
#accession        S60947
#molecule_type   DNA
#residues         1-265 ##label GAL
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772

REFERENCE
#authors          Boyer, J.; Fairhead, C.; Gallon, L.; Gallsion, F.; Michaux,
#submission       G.; Thierry, A.; Dujon, B.
#accession        S67104
#molecule_type   Submitted to the Protein Sequence Database, July 1996
#accession        S67113
#molecule_type   DNA
#residues         1-265 ##label BOY
#cross-references EMBL:L275128; NID:g1420509; PID:e252085; PID:g1420510
                  MIPS:Y0R220W

#experimental source strain S288C

REFERENCE
#authors          Gallsion, F.; Dujon, B.
#journal          Yeast (1996) 12:877-885
#title            Sequence and analysis of a 33 kb fragment from the right arm
                  of chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MUID:96437977
#accession        S71722
#status           Nucleic acid sequence not shown; translation not shown

GENERICS
#map_position     15R
SUMMARY           #length 265 #molecular_weight 29255 #checksum 6028

Query Match      3.6%; Score 113; DB 2; Length 265;
Best Local Similarity 35.2%; Pred. No. 2,77e-01;
Matches 25; Conservative 10; Mismatches 33; Indels 3; Gaps 3;

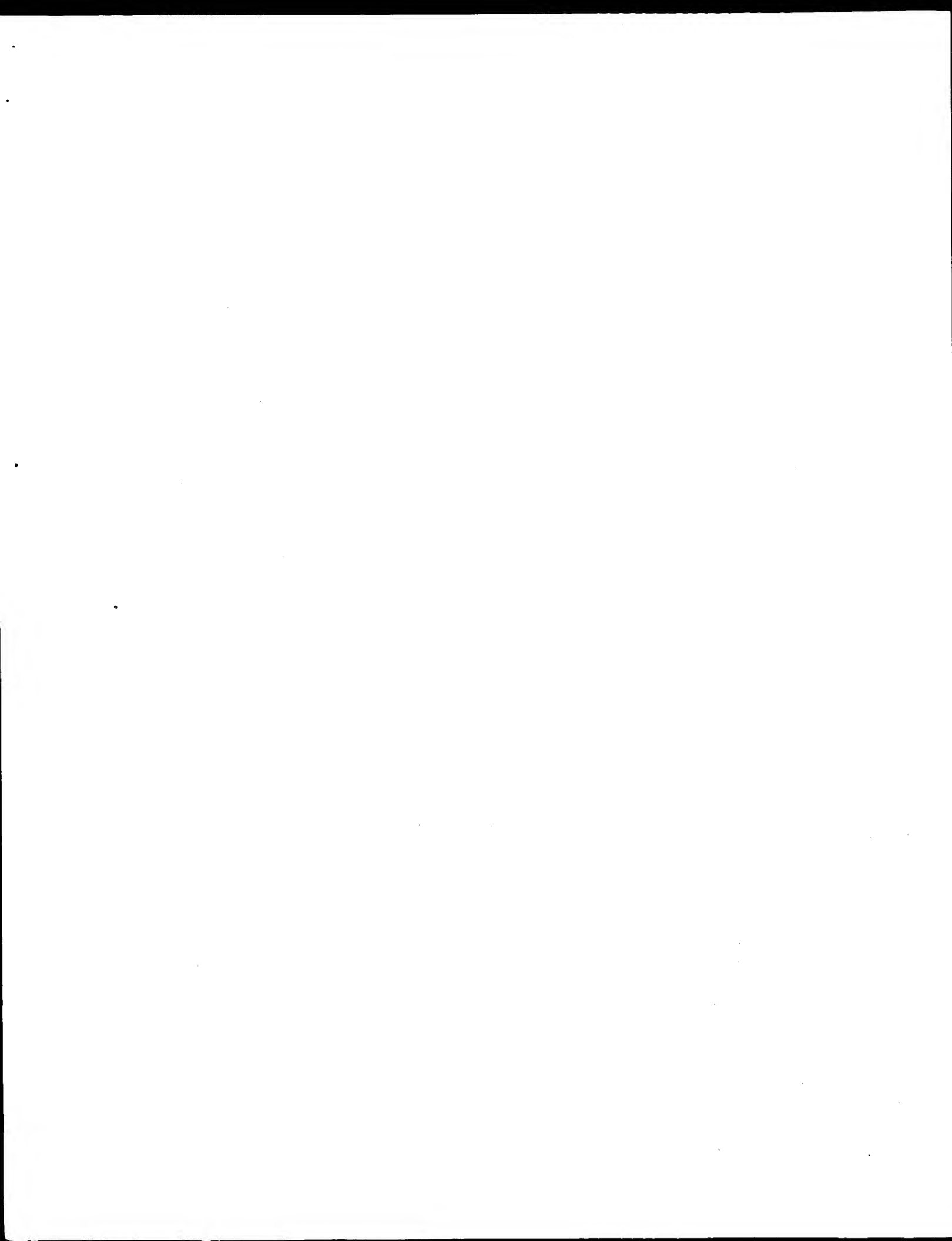
Db 118 LLEDRNRKPLSTINT-PCGVTVGVDSSTLNGGSSLSPPDKSSLEPTWLRKSTDSKPFYSQ 176
OY 43 FVXXERRKRSTSTSDIXYATXDSRLY-PGALLVDETLENNPTL-LAYDRAPRTYS 100
Db 177 EPLPRKSRSS 187
OY 101 XXLPLGLASSDS 111

RESULT 13
ENTRY          S03974 #type complete
TITLE          amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine
ALTERNATE_NAMES monoamine oxidase type A
ORGANISM       #format name Bos primigenius taurus #common name cattle
DATE           28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change
                26-Jan-1999

ACCESSIONS
REFERENCE
#authors          S03974
                  S03974
                  Powell, J.F.; Hsu, Y.P.P.; Weyler, W.; Chen, S.; Salach, J.;
                  Andrikopoulos, K.; Mallet, J.; Breakfield, X.O.
                  Biochem. J. (1989) 259:407-413
#journal          The primary structure of bovine monoamine oxidase type A.
#title            Comparison with peptide sequences of bovine monoamine
                  oxidase type B and other flavoenzymes.
#cross-references MUID:89246344
#accession        S03974
#status           not compared with conceptual translation
#molecule_type   mRNA

```

Db	449	IEY-YKRIYASDEGRI-KVFIIEELBELYKKT--GD-KRIFIGVKVKEGISTIVAVLAD	503
Oy	13	MNYDXKKLTTHQGESEIENRFKKEGNOXPFEVXXEKRRKRSITNTSDIXVATYDSRLYP	72
Db	504	GSIIPEELPEKAP-VVNIIRVFTDELFL	533
Oy	73	GALLVVDETXLENNPTLLAVDRAPMTYSXL	103



 W E S T
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mrsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Aug 30 16:10:12 1999; Maspar time 14.42 Seconds
 Tabular output not generated. 923.239 Million cell updates/sec

Title: >US-09-120-044-4
 Description: (1-471) from US09120044.pep
 Perfect Score: 3127
 Sequence: 1 MANKAVNDPILANVYDXKL.....TISNGTITPYQVEKQVEND 471

Scoring table: PAM 150
 Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot37
 1:swissprot

Statistics: Mean 51.211; Variance 95.336; scale 0.537

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	3116	99.6	470	1	TACY_STRPN	PNEUMOLYSIN (THIOL-ACT	0.00e+00
2	1556	49.8	500	1	TACY_CLOPE	LISTERIOLYSIN O PRECU	0.00e+00
3	1475	47.2	529	1	TACY_LISMO	LISTERIOLYSIN O PRECU	2.10e-290
4	1458	46.6	528	1	TACY_LISTIV	IVANOLYSIN PRECURSOR (1.43e-286
5	1416	45.3	501	1	TACY_PAEAL	ALVEOLYSIN PRECURSOR (4.14e-277
6	1414	45.2	530	1	TACY_LISSE	SEBILIGERIOLYSIN PRECU	1.17e-276
7	1379	44.1	571	1	TACY_STRPY	STREPTOLYSIN O PRECURS	8.80e-269
8	1380	44.1	574	1	TACY_STRPY	STREPTOLYSIN O PRECURS	5.24e-269
9	1363	43.6	574	1	TACY_STRCB	STREPTOLYSIN O PRECURS	3.49e-265
10	1298	41.5	485	1	TACY_BACCE	HEMOLYSIN PRECURSOR (F	1.40e-230
11	1112	3.6	527	1	MOFA_BOVIN	AMINE OXIDASE (FLAVIN	6.96e-02
12	108	3.5	462	1	MEPB_YEAST	MITOCHONDRIAL PROCESSI	2.34e-01
13	108	3.5	1228	1	SLAP_YEAST	S-LAYER PROTEIN PRECUR	2.34e-01
14	107	3.4	364	1	GCH2_POLE	GTP CYCLOHYDROLASE II	3.15e-01
15	105	3.4	776	1	ARCB_ECOLI	AEROBIC RESPIRATION CO	5.67e-01
16	105	3.4	780	1	CTPA_MYCLE	CATION-TRANSPORTING P-	1.35e+00
17	102	3.3	514	1	THD1_SALTY	THREONINE DEHYDRATASE	1.01e+00
18	103	3.3	722	1	IG42_YEAST	HYPOTHETICAL 78.8 KD P	1.35e+00
19	102	3.3	878	1	SYV_METJA	VALYL-TRNA SYNTHETASE	1.45e+00
20	99	3.2	279	1	YRR6_MYCCA	HYPOTHETICAL 33.0 KD P	1.79e+00
21	101	3.2	350	1	ARGI_SOYBN	ARGINASE (EC 3.5.3.1).	1.79e+00
22	100	3.2	427	1	GLYA_SYNY3	SERINE HYDROXYMETHYLTR	2.38e+00
23	100	3.2	502	1	SPEI_LYCES	ARGININE DECARBOXYLASE	2.38e+00

ALIGNMENTS

RESULT	ID	TACY_STRPN	STANDARD:	PRT:	470 AA.
AC	P11990;				
DT	01-OCT-1989 (REL. 12, CREATED)				
DT	01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN).				
GN	PLY.				
OS	STREPTOCOCCUS PNEUMONIAE.				
OC	BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;				
OC	STREPTOCOCCUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-WTCC 7466 SEROTYPE II;				
RX	MEDLINE: 87193109.				
RA	WALKER J.A., ALLEN R.L., PALMAGNE P., JOHNSON M.K., BOUINIS G.J.;				
RT	"Molecular cloning, characterization, and complete nucleotide				
RT	sequence of the gene for pneumolysin, the sulfhydryl-activated toxin				
RT	of Streptococcus pneumoniae.";				
RL	INFECT. IMMUN. 55:1184-1189(1987).				
CC	-!- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL				
CC	CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.				
CC	CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO				
CC	EUKARYOTIC CELL MEMBRANES.				
CC	-!- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.				
CC	This swiss-prot entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL: X52474; GA7404; -				
DR	EMBL: M17717; G153692; -				
DR	PIR: A28568; A28568.				
DR	PROSITE: PS00481; THIOL_CYTOLYSINS; 1.				
DR	PFAM: PF01289; Thiol_cytolysin; 1.				
DR	HSSP: P19995; 1PFO.				
KW	TOXIN; HEMOLYSIS; CYTOLYSIS.				
FT	INIT MET				
FT	ACT_SITE 427 427				
SO	SEQUENCE 470 AA; 52768 MW; D3F3A252 CRC32;				
FT	BINDING TO CHOLESTEROL (BY SIMILARITY).				
Query Match	99.6%; Score 3116; DB 1; Length 470;				
Best Local Similarity	94.5%; Pred. No. 0.00e+00;				

Matches 444; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

DB 1 ANKAVNDPILAMNDKRLKLLTHGSESIENRFIREGNLDPDEFVIEKKRSISTNTSDIS 60
 RT SHIMIZU T., KOBAYASHI T., BA-THIEN M., OHNANI K., HAYASHI H.;
 RT "Sequence analysis of flanking regions of the pfa gene of
 RT Clostridium perfringens: beta-galactosidase gene (pbg) is located in
 RT the 3'-flanking region.";
 RT MICROBIOL. IMMUNOL. 39:677-686(1995).
 DB 61 VTATNDSRLPGALLVDETLLENNPTLLAVDRAPMYTSIDLPGIASSDSFLOVEDPSNS 120
 RT IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
 RT IMMOLINE; 88004463.
 DB 62 VXAIXDSTLYGALLVDETLLENNPTLLAVDRAPMYTSXXLPGLASSDSFLOVEDPSNS 121
 RT OHNANO M., OHNO-IMASHITA Y., ANDO S.;
 RT "Role of the essential thiol group in the thiol-activated cytolysin
 RT from Clostridium perfringens.";
 RT EUR. J. BIOCHEM. 167:425-430(1987).
 DB 121 SVRGAVNDLAKKHQDYGQVNNVPAKQYKETAHSEQLKAFKGFSPETGNSLDIDFN 180
 RT X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RT FEEL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Crystallization and preliminary x-ray analysis of a thiol-activated
 RT cytolysin.";
 RT FEBS LETT. 397:290-292(1996).
 DB 122 SVRGAAXHLLAKKHQDYGQVNNVPAKQYKETAHSEQLKAFKGFSPETGNSLDIDFN 181
 RT X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RT FEEL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
 RT model of its membrane form.";
 RT CELL 89:685-692(1997).
 DB 181 SVHSGEKOIOIVNFKOITYTSDAVKKNPGDVFODITVEDLKRGISABERPLYISSVA 240
 RT FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 RT CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 RT CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 RT EUKARYOTIC CELL MEMBRANES.
 RT -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
 RT -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 DB 182 SVHSGEKOIOIVNFKOITYTSDAVKKNPGDVFODITVEDLKRGISABERPLYISSVA 241
 RT THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 RT between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RT the European Bioinformatics Institute. There are no restrictions on its
 RT use by non-profit institutions as long as its content is in no way
 RT modified and this statement is not removed. Usage by, and for commercial
 RT entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 RT or send an email to license@isb-sib.ch).
 DB 241 YGROYLTKLETTSSKDEVEAFALIKGVKVAPOTEWKQILDNTEVRAVILGGDPSSGAR 300
 RT EMBL; M81080; G144886; -;
 RT EMBL; M36704; G144884; -;
 RT EMBL; D49337; G1502275; -;
 RT PIR; B43577; B43577.
 DB 242 YXROYLTKLETTSSKDEVEAFALIKGVKVAPOTEWKQILDNTEVRAVILGGDPSSGAR 301
 RT PDB; 1PFO; 05-AUG-98.
 DB 301 VVTGKVDVDEDLIOGSEFTADHGPISYTTSTFLRNVAATPQNSDYEYETKVTAYRNG 360
 RT PROSITE; PS00481; THIOL-CYTOLYSINS; 1.
 DB 302 VVTGKVDVDEDLIOGSEFTADHGPISYTTSTFLRNVAATPQNSDYEYETKVTAYRNG 361
 RT PRAM; PF01289; Thiol cytolysin; 1.
 DB 361 DILDDHSGAAYAOYITTMDELSDYHOGKEVLTPKAMRNGODLTAHFTTSLPLKGNVRL 420
 RT TOXIN; HEMOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
 DB 362 DILDDHSGAAYAOYITTMDELSDYHOGKEVLTPKAMRNGODLTAHFTTSLPLKGNVRL 421
 RT STAIN-ATCC 13124;
 DB 421 SVKIRECTGLAMWNRVYEKTDLPVRRKRTISWGTLTYQVEDKYEVD 470
 RT TWETEN R.K.;
 DB 422 SVKIRECTGLAMWNRVYEKTDLPVRRKRTISWGTLTYQVEDKYEVD 471
 RT "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
 RT from Clostridium perfringens: significant homology with the genes for
 RT streptolysin O and pneumolysin";
 RT INFECT. IMMUN. 56:3235-3240(1988).
 DB RESULT 2 STANDARD: PRT: 500 AA.
 DB TACY_CLOPE
 DB AC P19995;
 DB DT 01-FEB-1991 (REL. 17, CREATED)
 DB DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DB DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DB DE PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).
 DB GN PFO OR PFOR OR PROA
 DB OS CLOSTRIDIUM PERFRINGENS.
 DB OG PLASMID PRT1B.
 DB OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
 DB CC CLOSTRIDIUM.
 DB RN [1]
 DB RP SEQUENCE FROM N.A.
 DB RX MEDLINE; 91099951.
 DB RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
 DB "An upstream regulatory sequence stimulates expression of the
 DB perfringolysin O gene of Clostridium perfringens.";
 DB INFECT. IMMUN. 59:137-142(1991).
 DB RN [2]
 DB RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.
 DB RC STRAIN-ATCC 13124;
 DB RX MEDLINE; 89032623.
 DB RA TWETEN R.K.;
 DB "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
 DB from Clostridium perfringens: significant homology with the genes for
 DB streptolysin O and pneumolysin";
 DB INFECT. IMMUN. 56:3235-3240(1988).
 DB RN [3]
 DB RP SEQUENCE OF 29-45 AND 305-312.
 DB RC STRAIN-DB6K;
 DB RX MEDLINE; 87076517.
 DB RA OHNO-IMASHITA Y., OHNANO M., MITSUI K., KAWASAKI H., ANDO S.;
 DB "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
 DB from Clostridium perfringens.";
 DB BIOCHEMISTRY 25:6048-6053(1986).
 DB RN [4]
 DB RP SEQUENCE OF 492-500 FROM N.A.

RC STRAIN-NCTC 8237;
 RX MEDLINE; 96123363.
 RA SHIMIZU T., KOBAYASHI T., BA-THIEN M., OHNANI K., HAYASHI H.;
 RT "Sequence analysis of flanking regions of the pfa gene of
 RT Clostridium perfringens: beta-galactosidase gene (pbg) is located in
 RT the 3'-flanking region.";
 RT MICROBIOL. IMMUNOL. 39:677-686(1995).
 DB 61 VTATNDSRLPGALLVDETLLENNPTLLAVDRAPMYTSIDLPGIASSDSFLOVEDPSNS 120
 RT IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
 RT IMMOLINE; 88004463.
 DB 62 VXAIXDSTLYGALLVDETLLENNPTLLAVDRAPMYTSXXLPGLASSDSFLOVEDPSNS 121
 RT OHNANO M., OHNO-IMASHITA Y., ANDO S.;
 RT "Role of the essential thiol group in the thiol-activated cytolysin
 RT from Clostridium perfringens.";
 RT EUR. J. BIOCHEM. 167:425-430(1987).
 DB 121 SVRGAVNDLAKKHQDYGQVNNVPAKQYKETAHSEQLKAFKGFSPETGNSLDIDFN 180
 RT X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RT FEEL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Crystallization and preliminary x-ray analysis of a thiol-activated
 RT cytolysin.";
 RT FEBS LETT. 397:290-292(1996).
 DB 122 SVRGAAXHLLAKKHQDYGQVNNVPAKQYKETAHSEQLKAFKGFSPETGNSLDIDFN 181
 RT X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RT FEEL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
 RT model of its membrane form.";
 RT CELL 89:685-692(1997).
 DB 181 SVHSGEKOIOIVNFKOITYTSDAVKKNPGDVFODITVEDLKRGISABERPLYISSVA 240
 RT FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 RT CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 RT CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 RT EUKARYOTIC CELL MEMBRANES.
 RT -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
 RT -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 DB 182 SVHSGEKOIOIVNFKOITYTSDAVKKNPGDVFODITVEDLKRGISABERPLYISSVA 241
 RT THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 RT between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RT the European Bioinformatics Institute. There are no restrictions on its
 RT use by non-profit institutions as long as its content is in no way
 RT modified and this statement is not removed. Usage by, and for commercial
 RT entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 RT or send an email to license@isb-sib.ch).
 DB 241 YGROYLTKLETTSSKDEVEAFALIKGVKVAPOTEWKQILDNTEVRAVILGGDPSSGAR 300
 RT EMBL; M81080; G144886; -;
 RT EMBL; M36704; G144884; -;
 RT EMBL; D49337; G1502275; -;
 RT PIR; B43577; B43577.
 DB 242 YXROYLTKLETTSSKDEVEAFALIKGVKVAPOTEWKQILDNTEVRAVILGGDPSSGAR 301
 RT PDB; 1PFO; 05-AUG-98.
 DB 301 VVTGKVDVDEDLIOGSEFTADHGPISYTTSTFLRNVAATPQNSDYEYETKVTAYRNG 360
 RT PROSITE; PS00481; THIOL-CYTOLYSINS; 1.
 DB 302 VVTGKVDVDEDLIOGSEFTADHGPISYTTSTFLRNVAATPQNSDYEYETKVTAYRNG 361
 RT PRAM; PF01289; Thiol cytolysin; 1.
 DB 361 DILDDHSGAAYAOYITTMDELSDYHOGKEVLTPKAMRNGODLTAHFTTSLPLKGNVRL 420
 RT TOXIN; HEMOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
 DB 362 DILDDHSGAAYAOYITTMDELSDYHOGKEVLTPKAMRNGODLTAHFTTSLPLKGNVRL 421
 RT STAIN-ATCC 13124;
 DB 421 SVKIRECTGLAMWNRVYEKTDLPVRRKRTISWGTLTYQVEDKYEVD 470
 RT TWETEN R.K.;
 DB 422 SVKIRECTGLAMWNRVYEKTDLPVRRKRTISWGTLTYQVEDKYEVD 471
 RT "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
 RT from Clostridium perfringens: significant homology with the genes for
 RT streptolysin O and pneumolysin";
 RT INFECT. IMMUN. 56:3235-3240(1988).
 DB RESULT 2 STANDARD: PRT: 500 AA.
 DB TACY_CLOPE
 DB AC P19995;
 DB DT 01-FEB-1991 (REL. 17, CREATED)
 DB DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DB DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DB DE PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).
 DB GN PFO OR PFOR OR PROA
 DB OS CLOSTRIDIUM PERFRINGENS.
 DB OG PLASMID PRT1B.
 DB OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
 DB CC CLOSTRIDIUM.
 DB RN [1]
 DB RP SEQUENCE FROM N.A.
 DB RX MEDLINE; 91099951.
 DB RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
 DB "An upstream regulatory sequence stimulates expression of the
 DB perfringolysin O gene of Clostridium perfringens.";
 DB INFECT. IMMUN. 59:137-142(1991).
 DB RN [2]
 DB RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.
 DB RC STRAIN-ATCC 13124;
 DB RX MEDLINE; 89032623.
 DB RA TWETEN R.K.;
 DB "Nucleotide sequence of the gene for perfringolysin O (theta-toxin)
 DB from Clostridium perfringens: significant homology with the genes for
 DB streptolysin O and pneumolysin";
 DB INFECT. IMMUN. 56:3235-3240(1988).
 DB RN [3]
 DB RP SEQUENCE OF 29-45 AND 305-312.
 DB RC STRAIN-DB6K;
 DB RX MEDLINE; 87076517.
 DB RA OHNO-IMASHITA Y., OHNANO M., MITSUI K., KAWASAKI H., ANDO S.;
 DB "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
 DB from Clostridium perfringens.";
 DB BIOCHEMISTRY 25:6048-6053(1986).
 DB RN [4]
 DB RP SEQUENCE OF 492-500 FROM N.A.

```

QY 123 VRAXXADLLAKHODYGVQVNNVPAKQYEXKTAHSEQLKVFSGPSEKXGNSLIDFNS 182
DB 214 VANNEKVMILAYKOJFYTSADLPKPNPSDLPDSDYTFNDLQKGSNAPPLMNSVAY 273
QY 183 VHSGEKXIQVNNKQIYITVSDAVKPNPGVDIVYEDLQKRGISARPLVYISXVAY 242
DB 274 GRITTYKLETTSSKDVQAFAKALIKNTDIKNSQYKDIYENSSFAVVLGGDAOEHNAY 333
QY 243 XROVYIKLETTSSXSEVEAFALIKGVAVAPQTEKKOLDINTYKAVILGGDPSSGARV 302
DB 334 VTKDFEIRKIVKDNATFEKKNAPYISTSVFLKNSVAANKTDYETSTESKCK 393
QY 303 VTKGVVMEEDLQESGRFTADHPGLPISTYTFSLRDVAVATFQNSDYETKVTAVRNGD 362
DB 394 INLHSGAVYAFVAMDEVSYDKEGNEVLTFRKTMGNOQDKTAKHSTYIPEANRNR 453
QY 363 LLLDHSGAVYAOYITWXLSELDHCKEVLTPKANDRNGDQLAHFTYISPLKGVNRNL 422
DB 454 IKARECTGLAMEMWMDYISEVDPPLTNINVSIMGTTLTP 493
QY 423 VKIRECTGLAMEMWRTVYEKTDLVKRRKRTISIMGTTLTP 462

RESULT 3
TACY LISTV STANDARD. PRT: 529 AA.
AC P31831.
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN HLY OR LISA OR HLYA.
OS BACTERIA MONOCYTOGENES.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD / SEROTYPE 1/2A.
RX MEDLINE: 89366684.
RA DOMANN E., CHAKRABORTY T.,
RT "Nucleotide sequence of the listeriolysin gene from a Listeria
RT monocytogenes serotype 1/2a strain."
RL NUCLEIC ACIDS RES. 17:6406-6406(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88153053.
RA MENGAUD J., VICENTE M.-F., CHENEVERT J., PEREIRA J.M., GEFEROY C.,
RT "Expression in Escherichia coli and sequence analysis of the
RT listeriolysin O determinant of Listeria monocytogenes."
RL INFECT. IMMUN. 56:766-772(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=12067.
RX MEDLINE: 92040062.
RA RASMUSSEN O.F., BECK T., OLSEN J.E., DONS L., ROSSEN L.,
RT "Listeria monocytogenes isolates can be classified into two major
RT types according to the sequence of the listeriolysin gene."
RT INFECT. IMMUN. 59:3943-3951(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=F2365, F5782 (SEROTYPE 4B), AND F4233, F6789 (SEROTYPE 1/2B);
RA VINES A., SWAMINATHAN B.;
RL SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DB EMBL: X15127; G44107; -
DR EMBL: M24199; G149653; -
DR EMBL: X60035; G44112; -
DR EMBL: U25452; G887870; -
DR EMBL: U25443; G887016; -
DR EMBL: U25446; G887028; ALT_INIT.
DR EMBL: U25449; G887864; ALT_INIT.
DR PIR: S05306; S05306.
DR PIR: A43505; A43505.
DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
DR HSP: P19995; lipo.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 529
FT ACT_SITE 484 484
FT VARIANT 35 35
FT VARIANT 39 39
FT VARIANT 438 438
FT VARIANT 523 523
SQ SEQUENCE 529 AA; 58688 MW; 269EA737 CRC32;

Query Match 47.2%; Score 1475; DB 1; Length 529;
Best local Similarity 40.88; Pred. No. 2,10e-290;
Matches 190; Conservative 115; Mismatches 160; Indels 1; Gaps 1;

DB 61 IDXYIQLDYNKNNVLYVGDATVNPVRKGYKDGNEYIVKCKKSNQNNADQVNA 120
QY 6 VNFILAMNDXKKLLTHGGESENEFKEGQLKREYVXXKRRKRSSTNSDIXVAT 65
DB 121 ISLTLPAGALVYKANSLELVNODVLPKRDSTLTSLDPLGNTQDNKIVYKATRSYNN 180
QY 66 XDRSLRPLGALLVDETXLNNPTLLAVDRPAPYKXPLGLASDPSPLQVEDPSSVRG 125
DB 181 AVNTLVERNNEXYAOAYPVSAKIDYDDDEMAYSQOLAKETAKAVNNSLNVNGAIS 240
QY 126 AXDDLAKHODYGOVN-NVPAKQYEXKTAHSEQLKVFSGPSEKXGNSLIDFNSVH 184
DB 241 EGRMOEVIKFOIYNNVNNNEPIRSEFGKAVTKEDQLAGVAENPPIYSSVAYGR 300
QY 185 SGKKXIQVNNKQIYITVSDAVKPNPGVDIVYEDLQKRGISARPLVYISXVAYXR 244
DB 301 OYTKLSTSHSTKYKAFDAVSKSVSGDELTNIRKNSFKAVIYGSAGKDEVOIID 360
QY 245 OYVLEKLETTSSXSEVEAFALIKGVAVAPQTEKKOLDINTYKAVILGGDPSSGARVY 304
DB 361 GNLDLRLDKKATFENETPGVPFAYTNTFLKNDLAVIKNSSEYITTKATTDGKIN 420
QY 305 GAVDWMEDLQESGRFTADHPGLPISTYTFSLRDVAVATFQNSDYETKVTAVRNGDL 364
DB 421 IDHSGIYAQENISWDEVNYPDEGNEIYQHKNSSENNKSLAHFTSSYIPEANRINNY 480
QY 365 IDHSGAVYAOYITWXLSELDHCKEVLTPKANDRNGDQLAHFTYISPLKGVNRNL 424
DB 481 AKRECTGLAMEMWRTVYEDRNPPIYKNNRISTWGTTLTPKYSNKKYDN 526
QY 425 IRECTGLAMEMWRTVYEKTDLVKRRKRTISIMGTTLTPQVEDKVEN 470

RESULT 4
TACY LISTV STANDARD. PRT: 528 AA.
AC P31831.
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE IVANOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).

```

```

IN ILO.
25 LISTERIA IVANOVII.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19119;
RX MEDLINE; 92182018.
RA HAAS A.; DUMBSKY M.; KREFT J.;
RT "Listeriolysin genes: complete sequence of llo from Listeria ivanovii
and of lse from Listeria seeligeri".
RL BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -1- FUNCTION: SULEFYLDRI-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
DR EMBL; X60461; E39053; ALT_INIT.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSP; P19995; lppo.
KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL
FT CHAIN 1 23
FT ACT_SITE 433 483 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 528 AA; 58542 MW; CFAE84AD CRC32;

Query Match 46.6% Score 1458; DB 1; Length 528;
Best Local Similarity 40.6%; Pred. No. 1,43e-286; Indels 1; Gaps 1;
Matches 189; Conservative 117; Mismatches 159;

Db 60 IDQYIGLDYDKNIIIVYDGEAVKNNVPKAGYEGNOYIVVEKKRSISNONNADIOYINS 119
QY 6 VDNFILAMNDXXXKLLTHGEGSIEINREXKKGNOLPBXFYXEXKKRSLSTNTSDIXXAT 65
Db 120 LASLTIPGALVKANSELVENQDYLPRKDSYVLSIDLPQGMVHNDNIIVYQNTKSIIND 179
QY 66 XDSRLTPGALLVYDERTLENNPTLLAVDRAPMYTXXLPGLASDSDFLOVEDSSNSVRG 125
Db 180 GVTNLYDVRMNNKSEEPNYSIAKIDYQOENAAVSESOVLARFGAFAKVAVNSLWNEGAIS 239
QY 126 AXXDLAKMHODYGOVN-NVPARQYIKKTAHSMEOCLKYVGFSDPEKXGNSLIDENSVH 184
Db 240 EGRVQEEVINFKQIYTVNVNEPTSPSRFPKSVTKENLOALGVNAENPPRATISSVAYGR 299
QY 185 SGEKKIQIVNNKQIYTVSVDAVKNPDQDQDITVEDLKRQGISASRPATYISIXVAYXR 244
Db 300 DIPKLTSSHSHTVKRAAFDIAEFKFGSVKGDTELENIIONASFVAALYVSGAKDEVELID 359
QY 245 QYVYKLTETTSXKXVEAEFAELLGVKAPOTEKMKQILDNTYKAAVILGGDPSSGARVYT 304
Db 360 GDLSEKLDILKOGANFPDKNKGVIATYTNLKNOLNAVNNSEYIETTSKAYSDEKIN 419
QY 305 GKVMWEDLLIOGSRFLADHGLTISYTSLLRDNVAVATPQNSDYETKVTAYRNGDLL 364
Db 420 LDHSGAVYARFNTWDEYSYDANGNEVEHKKKSENDKDLAHTTTSIYLPQGNRNINIH 479
QY 365 LDHSGAVYAOQYIYTWXELSYDHQGEVYLTPKAMPBNGDGLLAHFTTSLPLKGVNKNLSVK 424
Db 480 AKECTGGLAMWMAFVVDNRNLPIVKNRNVCIWGTTIYPAYSADYDYN 525
QY 425 IRECTGLAMWMAFVYETDILXVAKKRIISMGTTIYQVEDKAYDN 470

```

[illegible]

```

OY 250 LETTSKXSEVEAFALIKGVKAPQTEMKQIDNTYKAKVAILGDPSSGAVVTGKVM 309
DB 343 IQSVIKDNAQFSSKNPAPISYTSVFLKDNSIAVHNNTETKTEYSKGIKIDHSG 402
OY 310 VEDLIQESRRTADHPGLPISYTSFLRDVNAVAFQSTVDYELTKVAYARNGLDLHSG 369
DB 403 AYVAQFEYVDESDYDADGOEIVTRKSMWDMRDSAFESTELPLPPNKNKIFARECT 462
OY 370 AYVAQYITWELSYDHQKEVLTPKAMDRNGDLDLAHFTSLPLGNVNRNLSVKIRECT 429
DB 463 GLAEMWRTYVDEYNPLASDINVTGTLTP 495
OY 430 GLAEMWRTYVDEYNPLASDINVTGTLTP 462

RESULT 6
ID TACY_LISSE STANDARD: PRT: 530 AA.
AC P31830:
DT 01-FEB-1993 (REL. 26, CREATED)
DT 01-FEB-1993 (REL. 26, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE SEELIGERIOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ISO.
OS LISTERIA SEELIGERI.
OC BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
CC LISTERIA.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-SLCC:
RX MEDLINE: 92182018.
RA HAAS A., DUMBSKY M., KREFT J.;
RT "Listeriolysin genes: complete sequence of ilo from Listeria ivanovi1
and of ilo from Listeria seeligeri."
RL BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X60462; G44145; .
DR PIR: S22340; S22340.
DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DR PFM: PFO1289; Thiol_cytolysin; 1.
DR HSSP: P19995; IPO.
DR TOXIN: HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 25
FT CHAIN 26 530 SEELIGERIOLYSIN.
FT ACT_SITE 485 485 BINDING TO CHOLESTEROL (BY SIMILARITY).
SO SEQUENCE 530 AA; 59181 MW; 68A457A5 CRC32;

Query Match 45.2%; Score 1414; DB 1; Length 530;
Best Local Similarity 40.6%; Pred. No. 1.17e-276;
Matches 189; Conservative 109; Mismatches 167; Indels 1; Gaps 1;

DB 62 INKYIWMGLNYSKNSLTVYOGSAVNPVPPKGYKDGSEIYVEKKKGGINONNADISVINA 121
OY 6 VNDFLANNTYDXXKLLTHQGESIENREXKEGNOPLXEFVXXEKRRKSLSTNTSDIYVXAT 65
DB 122 ISSLTGALVKNARELVENQPNVLPVKRDLSTLVLPMTKKDKNFIYKNTFSNVAN 181
OY 66 XDSRLYPCALLVDETAKENPFTLLAVDRAPMTYSXALPGLASSDSFLOVEDPSNSVAG 125
DB 182 AVNTLVEMWNRKYSKAIPIINAKIDYSDENAYSESQLIAGFGAFPAVNNLSLVNFEAS 241

```

```

OY 126 AXDXLLAKMHQDYGQVN-NVPAKQYKEXKTAHSMEOIAKVFSGDEPKXGNSLDIDFNSVH 184
DB 242 DKGQOEVEISFKQIYYNNINNEPTSPKFFGSGYKREQLDALGVANENPVAISSVAYGR 301
OY 185 SGEXKXIOIVAKKQIYTVSDAVKKNPGDVEFQTVVEDLKQGIASERPLVYISXVAYXR 244
DB 302 QVYVKKSSSSHSNKVKTAFEAAMGSKVAGDELNTIINKSEFKYIVYIGSKAKEVELID 361
OY 245 QVYKLETTTSXSEVDAEAFALIKGVKAPQTEMKQIDNTYKAKVAILGDPSSGARVYT 304
DB 362 GNIGELRLDKRSTDRNPGVPISTYTNELKNDLAVVKNNSIEYETSKSYTDGKIN 421
OY 305 GKVDWVEDLIQESRRTADHPGLPISYTSFLRDVNAVAFQSTVDYELTKVAYARNGLDL 364
DB 422 IDHSGYVAQFNISMDVEYDENGNEIKVKKMGKNGENYSKLAHFTSSLYLGNARNINITY 481
OY 365 LDHSGAYVAQYITWELSYDHQKEVLTPKAMDRNGDLDLAHFTSLPLGNVNRNLSVK 424
DB 482 ARCTGLFEMWRTYVDEYNPLASDINVTGTLTPRHSNVN 527
OY 425 IRCTGLAEMWRTYVDEYNPLASDINVTGTLTPRHSNVN 470

RESULT 7
ID TACY_STREPY STANDARD: PRT: 571 AA.
AC P21131:
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
CC STREPTOCOCCUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 88057628.
RA KEROE M.A., MILLER L., WALKER J.A., BOULINOIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural
RT homologies between SLO and other membrane-damaging, thiol-activated
RT toxins."
RL INFECT. IMMUN. 55:3228-3232(1987).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M16638; G15811; .
DR PIR: A43507; A43507.
DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DR PFM: PFO1289; Thiol_cytolysin; 1.
DR HSSP: P19995; IPO.
DR TOXIN: HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMAID.
FT SIGNAL 1 33
FT CHAIN 34 571 STREPTOLYSIN O.
FT ACT_SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SO SEQUENCE 571 AA; 63638 MW; 33124E54 CRC32;

Query Match 44.1%; Score 1379; DB 1; Length 571;
Best Local Similarity 40.7%; Pred. No. 8.80e-269;
Matches 186; Conservative 100; Mismatches 170; Indels 1; Gaps 1;

DB 109 INDKIYSLNNEVLAKNGETIENFVPEKGVKADKFIYERKKKNINTPTVDISIDS 168

```



```

CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEMINATION OF BIOGENIC AND
CC XENOBIOTIC AMINES AND HAS IMPORTANT FUNCTIONS IN THE METABOLISM OF
CC NEUROACTIVE AND VASOACTIVE AMINES IN THE CENTRAL NERVOUS SYSTEM
CC AND PERIPHERAL TISSUES. MAO-A PREFERENTIALLY OXIDIZES BIOGENIC
CC AMINES SUCH AS 5-HYDROXYTRYPTAMINE (5-HT), NOREPINEPHRINE AND
CC EPINEPHRINE.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: MONOMER, HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X15609; G524; -.
CC DR EMBL: X15609; G525; ALT_INIT.
CC DR PIR: S03974; S03974.
CC KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; TRANSMEMBRANE; MITOCHONDRION;
CC NEUROTRANSMITTER DEGRADATION.
CC NP_BIND 14 70 FAD (ADP PART) (POTENTIAL).
CC FT BINDING 406 406 FAD (BY SIMILARITY).
CC FT TRANSMEM 498 518 POTENTIAL.
CC SQ SEQUENCE 527 AA; 59800 MW; 724E4396 CRC32;

Query Match 3.6%; Score 112; DB 1; Length 527;
Best Local Similarity 18.7%; Pred. No. 6,96e-02;
Matches 28; Conservative 42; Mismatches 74; Indels 6; Gaps 5;

Db 1 MESLOKTSIADGOMEDVIVIGGSGISGSAK-LIAEHVAVLVLEARERGGTIVYRNH 59
QY 250 LETTSSXKEVEAEALIGVAVAPOTEWKQILLDNTRYKAVILGGPSSGAVVVGKADM 309
Db 60 V-YVVGAGVAVPTQNRILRLSKGLFETKYVNERLVHVKGRTYFRRAPFPVNP 118
QY 310 VEDLIQGSFTDHPGLPISTYTSF-LADNVATFQNSTDIYETVAVYRNGDLLDHS 368
Db 119 IAYL-DYNNILRTM--DNNKKEIPADAPWE 145
QY 369 GAVVAQYIITWXLXSLYDHQKEVLTPKAMD 398

RESULT 12
ID MPPB_YEAST STANDARD: PRT; 462 AA.
AC P10507;
DT 01-JUL-1989 (REL. 11, CREATED)
DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DE MITOCHONDRIAL PROCESSING PEPTIDASE BETA SUBUNIT PRECURSOR
DE (EC 3.4.24.64) (BETA-MP) (PEP).
GN MASI OR MIF1 OR YLR163C OR L9632.10.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88312592.
RA WITTE C., JENSEN R.E., YAFFE M.P., SCHATZ G.;
RT "MAS1, a gene essential for yeast mitochondrial assembly, encodes a
RT subunit of the mitochondrial processing protease."
RL EMBL J. 7:1439-1447(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-S288C / AB972;
RC JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,

```

```

RA FAVELLO A., FULTON L., GATTUNG S., GRECO T., KIRSTEN J.,
RA KOCABA T., HALSWORTH K., HAWKINS J., HILLER L., JIER M.,
RA JOHNSTON D., JOHNSTON L., LANGSTON Y., LATREILLE P., LE T.,
RA NARDIS E., MENZES S., MILLER N., NHAN M., PAULEY A., PELUSO D.,
RA RIEKEN L., RILES L., TAICH A., TREWASKIS E., YAGNATI D.,
RA WILCOX L., WOLDMAN P., VAUDIN M., WILSON R., WATERSTON R.;
RA SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RL [3]
RN SEQUENCE OF 21-32.
RP MEDLINE; 91177897.
RX YANG M., GELI V., OPLIGER W., SUDA K., JAMES P., SCHATZ G.;
RT "The MAS-encoded processing protease of yeast mitochondria.
RT Interaction of the purified enzyme with signal peptides and a
RT purified precursor protein."
RL J. BIOL. CHEM. 266:6416-6423(1991).
CC -1- FUNCTION: THE MITOCHONDRIAL PROCESSING PROTEASE CLEAVES
CC PRESEQUENCES FROM MITOCHONDRIAL PROTEIN PRECURSORS.
CC -1- CATALYTIC ACTIVITY: RELEASE OF N-TERMINAL TRANSIT PEPTIDES FROM
CC PRECURSOR PROTEINS IMPORTED INTO THE MITOCHONDRION, TYPICALLY WITH
CC ARG IN POSITION P2.
CC -1- COFACTOR: REQUIRES DIVALENT CATIONS FOR ACTIVITY.
CC -1- SUBUNIT: HETERODIMER OF ALPHA AND BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M16; ALSO KNOWN AS THE
CC INSULINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X07649; G3887; -.
CC DR EMBL: U01921; G1234852; -.
CC DR PIR: S00552; S00552.
CC DR PIR: A38734; A38734.
CC SGD; L0001026; MASI.
CC DR PROSITE; PS00143; INSULINASE; 1.
CC DR PRAM; PF00675; Insulinase; 1.
CC KM HYDROLASE; METALLOPROTEASE; ZINC; MITOCHONDRION; TRANSIT PEPTIDE.
CC FT TRANSIT 1 20 MITOCHONDRIAL PROCESSING PROTEASE
CC FT CHAIN 21 462 BETA SUBUNIT
CC FT METAL 70 70 ZINC (BY SIMILARITY).
CC FT ACT_SITE 73 73 ZINC (BY SIMILARITY).
CC FT METAL 150 150 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 462 AA; 51083 MW; 438D1777 CRC32;

Query Match 3.5%; Score 108; DB 1; Length 462;
Best Local Similarity 21.9%; Pred. No. 2,34e-01;
Matches 33; Conservative 45; Mismatches 63; Indels 10; Gaps 9;

Db 65 NNGTAHLEHLAEK-GTQ-NRSOGIELEIET-IGSHLNAVTSRENTVYAKSLQDIPK 121
QY 151 EKXTRAHMEQKAKFGSDFEKXGNSLIDFNVSSEKXIQIYNKQIITYTSV-DAYAN 209
Db 122 AVDLISDLIKSVLDNSAIRERDVILRESEEDKMYDEVDHLHEITIKDQPLGRITL 181
QY 210 PGDFQDTIVIEDLKQNGISAERPIYVI-SX-VA--YXRQVYKL-ETTSXSKVEVAARE 264
Db 182 GPIRKISITRIDIKDYITTKYKGRMYLAG 212
QY 265 ALIKGVAVAPOTEWKQ-ILLDNTRYKAVILGG 294

RESULT 13
ID SLAP_BACST STANDARD: PRT; 1228 AA.
AC P35825;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)

```


[illegible]

RT "Insights into multistep phosphorelay from the crystal structure of
RT the C-terminal HPT domain of ArcB";
RL CELL 88:717-723(1997).
TA

RP X-RAY CRYSTALLOGRAPHY (1.57 ANGSTROMS) OF 657-774.

RL SUBMITTED (APR-1998) TO THE PDB DATA BANK.

CC - FUNCTION: MEMBER OF THE TWO COMPONENT REGULATORY SYSTEM, WHICH
CC SENSOR-REGULATOR PROTEIN FOR ANAEROBIC REPRESSION OF THE ARC

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC MODULON. ACTIVATES ARCA BI PHOSPHORYLATION.

CC (PROBABLE).
CC -1- PTM: ACTIVATION PROBABLY REQUIRES A TRANSFER OF A PHOSPHATE GROUP

CC BETWEEN A HIS IN THE TRANSMITTER DOMAIN AND A ASP OF THE RECEIVER
CC DOMAIN (BY STIMULABILITY)

CC -!- SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE

CC -1- SIMILARITY: IN ITS C-TERMINUS, TO THE N-TERMINAL REGION OF THE
 -1- SIMILARITY: IN ITS C-TERMINUS, TO THE N-TERMINAL REGION OF THE
 -1- SIMILARITY: IN ITS C-TERMINUS, TO THE N-TERMINAL REGION OF THE

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X53315; G40951; -

DR EMBL; AE000400; G1789603; -
DR EMBL; G1800000; G1800000; -

DR PIR; SII/54, NOECHAN.
DR PDB; 1A0B; 18-MAR-98.

DR PDB; 2AUB; 17-JUN-98
DR ECGENE; EG10062; AF

DR	PEAM;	resp
DR	PEAM;	signa

DR PEAM; PF00989; PAS;
RECEIVED BY TRANSMISSION

KW TRANSMEMBRANE; INNER ME
 1 33

23	TRANSMEM	50	P
23	TRANSMEM	50	P

11	DOMAIN	58	77
22	TRANSMEM	58	77

COMAIN	18	110
PT	269	515

FT	DOMAIN	516	//6	RECEIVER DOMAIN (POTENTIAL).
FT	MOD_RES	292		PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FILE	LINE	TEXT
FILE	574	MOD_RES
FILE	574	CONFLICT
FILE	468	468
FILE	468	G -> GTG (IN REF. 1) ..
FILE	468	PHOSPHORYLATION (BI S

50 SEQUENCE 776 AA; 87824 MW; 45AB596D CRC32;

Query Match	3.48;	Score 105;	DB 1;	Length
Best Match	19.69;	Pred NO	5	67e-01.

Matches 37; Conservative 56; Mismatches 97; In

Db 246 KVPYDVGKRGHGMFGGRDITERKRYQDALERASRD-KTFISTIS

124 RGAXXDLAKWHQDYGOVNNVPAKXQYEKXTAHSMEQLKVKFGSDF-

Db 304 LSRILLDTeltaEQEKYlKtIHVSaVt-LGNIENDIIDMDKMERKRV

00 183 VHSGEKXIOTVNXKOIYY-TVSVDANKPGEVDFODVTVEDLKORGI

Db 363 ADI ENI SAI 0A00KGI RENI EPTI PI PHOV ITDGTBRLROIWNLI

```
0- | : | : ::|::: | | :: | :|| |
241 xvydov -tzi etmcysevevnafeaf tkcvkvaBOTFEWKOITDNTX
```

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

[illegible]

QY 300 ANVYIGVEMVEEELIYBOS JIC

1

(TM)

Distribution rights by Oxford Molecular Ltd

Tabular output not generated.

Sequence:

Scoring table:

Searched:

Database:

13:sp_vertebrate 14:sp_virus

Statistics:

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1731	55.4	497	2	055996	SULFASIN.	0.00e+00
2	1726	55.2	497	2	085102	HEMOLYSIN.	0.00e+00
3	1243	39.8	534	2	C31241	POLYSIN.	2.24e+23
4	125	4.0	50	2	048772	LISTERIOLYSIN O (FRAGM	1.01e-03
5	119	3.8	50	2	048773	LISTERIOLYSIN O (FRAGM	7.42e-03
6	113	3.6	265	3	012044	CHROMOSOME XV READING	5.12e-02
7	113	3.6	450	3	068518	HYPOHETICAL 4.7 KD P	5.12e-02
8	112	3.6	744	2	067108	DNA GYRASE A SUBUNIT.	7.01e-02
9	114	3.6	1361	2	087083	S-LAYER PROTEIN.	3.73e-02
10	114	3.6	1361	2	C30524	S-LAYER PROTEIN.	3.73e-02
11	109	3.5	218	1	038691	HYPOHETICAL PROTEIN M	1.79e-01
12	110	3.5	550	2	0686015	GLUCOSE-6-PHOSPHATE IS	1.31e-01
13	109	3.5	710	2	085906	TRANSPONSON PROTEIN B.	1.79e-01
14	108	3.5	814	10	065715	HYPOHETICAL 91.9 KD P	2.43e-01
15	106	3.4	426	2	067578	HYPOHETICAL 91.9 KD P	2.43e-01
16	103	3.3	446	2	034363	FLAGELLAR HOOK ASSOCIA	4.47e-01
17	100	3.2	446	2	034363	PUTATIVE L-AMINO ACID	1.10e+00
18	99	3.2	147	10	049900	MTN1A (FRAGMENT).	2.64e+00
19	99	3.2	182	1	058072	187AA LONG HYPOHETICA	3.52e+00
20	100	3.2	231	11	035813	KEATININ 14 (FRAGMENT).	3.52e+00
	100	3.2	287	1	059586	287AA LONG HYPOHETICA	2.64e+00

21	99	3.2	346	13	093404	PROLACTIN RECEPTOR (PR	3.52e+00
22	100	3.2	349	1	058654	HYPOTHETICAL PROTEIN M	2.64e+00
23	100	3.2	362	2	050434	AMINOTRANSFERASE.	2.64e+00
24	99	3.2	366	1	029257	ACC TRANSPOSER, ATP-B	3.52e+00
25	101	3.2	453	3	788803	YEAST.	1.97e+00
26	101	3.2	501	1	028947	CONSERVED HYPOTHETICAL	1.97e+00
27	99	3.2	580	2	033559	TRANSUDOR-LIKE PROTEI	3.52e+00
28	99	3.2	821	5	093378	C44H4.7 PROTEIN.	3.52e+00
29	99	3.2	877	3	006593	SMILAIR TO S. CEREVISI	3.52e+00
30	99	3.2	903	3	087330	DYNAMIN-RELATED PROTEI	3.52e+00
31	99	3.2	1035	2	025887	CATION EFFLUX SYSTEM P	3.52e+00
32	99	3.2	1329	5	076366	C45G7.6 PROTEIN.	3.52e+00
33	100	3.2	1792	5	094882	MOSTIN V (DILUTE CLASS	2.64e+00
34	100	3.2	5027	11	062442	MULTIPLY SUPPRESSOR (2.64e+00
35	96	3.1	281	3	051923	DEOXYRIDINE TRIPHOSH	6.22e+00
36	97	3.1	283	5	058761	HYPOTHETICAL PROTEIN M	6.22e+00
37	97	3.1	291	1	058761	TRANSCRIPTIONAL TERMIN	4.69e+00
38	98	3.1	436	2	067031	7-KETO-8-AMINODELARGON	8.24e+00
39	96	3.1	437	2	074770	SMILAIRTY TO C2H2-TYP	8.24e+00
40	96	3.1	498	5	068164	SENSOR PROTEIN BVS.	8.24e+00
41	96	3.1	601	2	090994	190 KD TENNASCIN PRECR	4.69e+00
42	98	3.1	1532	13	090995	200 KD TENNASCIN PRECR	4.69e+00
43	98	3.1	1714	13	090995	CITOSOLIC 200KD PRECU	4.69e+00
44	98	3.1	1810	13	090824	H50114.2 PROTEIN.	6.22e+00
45	97	3.1	2160	5	017709		

ALIGNMENTS

ID	RESULT	1
AC	055996:	PRELIMINARY;
AC	055996:	055997:
DT	01-NOV-1986	(TREMBLREL. 01, CREATED)
DT	01-JAN-1989	(TREMBLREL. 09, LAST SEQUENCE UPDATE)
DT	01-JAN-1999	(TREMBLREL. 09, LAST ANNOTATION UPDATE)
DE	SUTLISIN.	
GN	SIX.	
OS	STREPTOCOCCUS SUT.	
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE	
OC	STREPTOCOCCUS.	
RN	[1]	
RN	SEQUENCE OF 1-71 FROM N.A.	
RC	STRAIN=PI/77;	
RA	SEEBER R.P.A.M., KENTER T., DE HAAN W.A.M., JACOBS A.A.C.;	
RT	"Characterization of the gene encoding sul1ysin from <i>Streptococcus</i>	
RT	<i>SUT.</i> and expression in field strains."	
RL	SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.	
DO	EMBL; Z36907: E1334304;	
DO	SEQUENCE 497 AA: 54850 MW; D78AE8BF CRC32:	
		PRT; 497 AA.

Query Match 55.4%; Score 1731; DB 2; Length 497;

Matches	227;	Conservative	105;	Mismatches	134;	Indels	2;	Gaps	2.
---------	------	--------------	------	------------	------	--------	----	------	----

Db	29	SKODINOCYQSLTYPEROILLNEBEYIEDNPATTCMLENGRIYVLRREKKNITNNSADIA	88
Qy	2	ANAVADPFLANNYDXXKLLTHGSESTIENEKXEQDLPXEEVXXERKRKSLSTNTSDIX	61
Db	89	VIDAKANIYPGALLRADONLLDNNPILIS IARGDITLTLNPLGLANDSHTVNSPTRS	148
Qy	62	VXATXDSRFLYPCALLVDETLENNPITLAVDRAPTTYSXKLPGLAASDSLQYEDPSNS	121
Db	149	TYRTGYNNLLSMMNTTAYGEFNTQAELOYDETFMAYSMGLTKRGTSEFKIAPVDINF	208
Qy	122	SVRGAXXDILAWHODY-GOVNNPAPXQXQEKXTAISMGLTKVKGSPFEKXGNSLIDF	180
Db	209	DAVNSGKROVIYNEKOIYYVSVDEPESPSKFLAEGTVEBTKRNGTIDEVPPVYSSV	268
Qy	181	NSVHSEKXKIQVKNKQIYYTVSDAVKKNNGDFQDQTYVIEDLKORGISNAERPLVYISXV	240
Db	269	SYGRSMFIKLETSRSRTQYQAARAKAIKVIDISGNAEYODILIKNTSEFATYEGGDAGSAA	328
Qy	241	ATXROVYIKLETTSSKSEVEAEFAALIKGVKAAPOCEMKOILIDNNTYKAVAVILGGDPSSGA	300


```

RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE; 98196666.
RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEER R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RL aeolicus."
RL NATURE 392:353-358(1998).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RA DECKERT G., WARREN P.V., GASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBEER R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000716; G2983485;
SQ SEQUENCE 744 AA; 84080 MW; 93795328 CRC32;

Query Match
Best Local Similarity 24.2%; Pred. No. 7.01e-02;
Matches 22; Conservative 26; Mismatches 37; Indels 6; Gaps 5;

Db 449 IEY-YKKIVASGERT-KVFIEFEELVKKY--GD-KRRTIGGVKKKESSIVAVLQD 503
OY 13 MNYDXXKLLTHOGESIEHFKKGNQDPXFEVXXERRKRSLSSTSDIXXATYDSRLYP 72
DB 504 GSIIPEELPLEKAP-VVNIIRVPTGGLFL 533
OY 73 GALLVVDERTXLENNPTLLAVDRAPMTYSXL 103

RESULT
9
ID 087083 PRELIMINARY; PRT; 1361 AA.
AC 087083;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE S-LAYER PROTEIN.
OC CAMPYLOBACTER RECTUS.
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
OC CAMPYLOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33238;
RX MEDLINE; 98442421.
RA MIYAMOTO M., MAEDA H., KITANAKA M., KOKEGUCHI S., TAKASHIBA S.,
RA MURAYAMA Y.;
RT "The S-layer protein from Campylobacter rectus; sequence
RT determination and function of the recombinant protein.";
DR EMBL; AB001876; D1034498;
KW S-LAYER.
SQ SEQUENCE 1361 AA; 144904 MW; 55A9D736 CRC32;

Query Match
Best Local Similarity 19.9%; Pred. No. 3.73e-02;
Matches 37; Conservative 50; Mismatches 89; Indels 10; Gaps 9;

Db 178 VDMHSNPEHGRKAYELTTTNDATANVFNAPMK-HNPGGTDRIIMTLOSSDKLTGDSRHD 236
OY 115 VEDPSNSSVRGAXXDLLAKWHODYGOVNNVPAKXOYEKKTAHSM-QLKVFSGDFEXKG 173
DB 237 NTLNVEEGGANDGDPSTRTPLTINIONINIEVTGYNTLDL-RDSNDVEKINIRHTK 295
OY 174 NSLIDIFNSVHS--GEKXIQIVNKKQI-YYTVSDAVKPNPGDVFODTVVDLKKORGISA 230
DB 296 EAGNKNVNSIG-QKLYGRLANVAK-KDIDVKEH-KKGVLSGEDSKSNVLEAVEKS 352
OY 231 ERPLVY-ISKVAXXROYVTKLETTSKSXKEVAEPALIKGVAVAPOTEMKQILDNTYKA 289
DB 353 LSITSD 358

```

```

OY 290 VILGSD 295

RESULT
10
ID 030524 PRELIMINARY; PRT; 1361 AA.
AC 030524;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE S-LAYER PROTEIN.
GN CRS.
OS CAMPYLOBACTER RECTUS.
OC BACTERIA; PROTEOBACTERIA; EPSILON SUBDIVISION; CAMPYLOBACTER GROUP;
OC CAMPYLOBACTER.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-314;
RX MEDLINE; 98187925.
RA WANG B., KRAIG E., KOLODRUBETZ D.;
RT "A new member of the S-layer protein family: characterization of the
RT crs gene from Campylobacter rectus.";
RL INFECT. IMMUN. 66:1521-1526(1998).
DR EMBL; AF010143; G2459961;
SQ SEQUENCE 1361 AA; 144385 MW; CAFE081F CRC32;

Query Match
Best Local Similarity 19.9%; Pred. No. 3.73e-02;
Matches 37; Conservative 50; Mismatches 89; Indels 10; Gaps 9;

Db 178 VDMHSNPEHGRKAYELTTTNDATANVFNAPMK-HNPGGTDRIIMTLOSSDKLTGDSRHD 236
OY 115 VEDPSNSSVRGAXXDLLAKWHODYGOVNNVPAKXOYEKKTAHSM-QLKVFSGDFEXKG 173
DB 237 NTLNVEEGGANDGDPSTRTPLTINIONINIEVTGYNTLDL-RDSNDVEKINIRHTK 295
OY 174 NSLIDIFNSVHS--GEKXIQIVNKKQI-YYTVSDAVKPNPGDVFODTVVDLKKORGISA 230
DB 296 EAGNKNVNSIG-QKLYGRLANVAK-KDIDVKEH-KKGVLSGEDSKSNVLEAVEKS 352
OY 231 ERPLVY-ISKVAXXROYVTKLETTSKSXKEVAEPALIKGVAVAPOTEMKQILDNTYKA 289
DB 353 LSITSD 358
OY 290 VILGSD 295

RESULT
11
ID 058691 PRELIMINARY; PRT; 218 AA.
AC 058691;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN MJ1295.
GN MJ1295.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
OC METHANOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERRIDGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEER R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
RA SCOTT J.L., GEOGHEGAN N.S.M., WEIDMAN J.F., FUHRMAN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KATNE B.P., BORODOVSKY M.,
RA KLEIN H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RT "complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL SCIENCE 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG, TO A.FLIGIDUS AF1550 AND
CC M.THERMOPHILICUM MTH576.

```

```

DR EMBL: U67570; G1591933; -.
KW HYPOTHETICAL PROTEIN.
SEQUENCE 218 AA; 24666 MW; CFC99AB9 CRC32;

Query Match
Best Local Similarity 3.5%; Score 109; DB 1; Length 218;
Matches 19; Conservative 12; Mismatches 29; Indels 2; Gaps 2;

Db 86 KRAKDIAGAEIVVH-GEFVPEVEEKTNYASISDEDVILAFPGFDKETAENKENDIF 144
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 171 KXGNSLDIDFNSVHSGEKXIVXNKQIYTVSDAVKNPGDV-FODTVYEDLKRGIS 229
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 145 VE 146
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 230 AE 231

RESULT 12
ID 088015 PRELIMINARY; PRT; 550 AA.
AC 088015.
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE GLUCOSE-6-PHOSPHATE ISOMERASE.
GN PGI.
OS STREPTOMYCES COELICOLOR.
OC BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC ACTINOMYCETALES; STREPTOMYCINAE; STREPTOMYCETACEAE; STREPTOMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA SEGER K.J., HARRIS D.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.;
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA MEDLINE; 97000351.
RA REBENACH M., KIESER H.M., DENAPATTE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.;
RT "A set of ordered contigs and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MCL. MICROBIOL. 21:77-96(1996).
DR EMBL: AL031107; E1311951; -.
KW ISOMERASE.
SO SEQUENCE 550 AA; 60424 MW; 028FD1F6 CRC32;

Query Match
Best Local Similarity 3.5%; Score 110; DB 2; Length 550;
Matches 49; Conservative 53; Mismatches 104; Indels 17; Gaps 15;

Db 245 ALSTNAEKVADFCDITNMFEFMDWVGKRSFDSAIGLSMIAIGDPRFEMLDGFRIVD 304
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 52 SLSTNSIDIVXATXDSRLPGALLVVDVETXLENNPFL-LAYVRAPMTYSXLPGL-ASS 109
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 305 EHRFNAEPANAPL--LLGLGWYGDFLGAOSHAVLPYSHLSLFTAY-LOQLMESN 360
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 110 DSFLQVEDPNSSVRGAXXDLAKWHODY-G-OVNNVPAKXQI-EKXTAHSMQLKVKF- 165
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 361 GKSVDREGNPNVQOTGPPVWG-TP-GT-NGQNAVYQOLIHGOTKLIPADFLGARPYDELIS 417
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 166 GSPFEKXGSLDIDFNSVHSGEKXIVXNKQIYTVSDAVK-NGGDVFODTIYVEDLK 224
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 418 E-ELKAQHDLMANFA-OTQA-LAFGKTPDEVRAAGVPEELV 457
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 225 QRGISARPLVYISVAVYKQVTLKLETTYSXSXEVAFAEAL 267
| : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 13

```

```

ID 085906 PRELIMINARY; PRT; 710 AA.
AC 085906;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TRANSFOSONAS PROTEIN B.
GN TNPB.
OS SPHINGOMONAS AROMATICIVORANS.
OG PLASMID PN1.
OC BACTERIA; PROTOBACTERIA; ALPHA SUBDIVISION; ZYMONOMAS GROUP.
OC SPHINGOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-F199;
RA ROMINE M.F., STILLMELT L.C., WONG K.-K., THURSTON S.J., SISK E.C.,
RA SENSEN C.W., GASTERLAND T., SAFER J.D., FREDRICKSON J.K.;
RT "Complete sequence of a 184 kb catabolic plasmid from Spingomonas
RT aromaticivorans strain F199.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF079317; G3378338; -.
KW PLASMID.
SO SEQUENCE 710 AA; 81079 MW; 41B48670 CRC32;

Query Match
Best Local Similarity 3.5%; Score 109; DB 2; Length 710;
Matches 14; Conservative 20; Mismatches 25; Indels 3; Gaps 3;

Db 139 NHHNDVLECYAAWYGANGRR-ROSOT-ROMRGKITYSEFSPVQLGDIRFAIT 196
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 366 DHGAVVAQYIYTWELSDHSGKEVLTIPKANDRGQDILAFHTTS-IPLGKWNFLSVK 424
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 197 NR 198
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 425 IR 426

RESULT 14
ID 065715 PRELIMINARY; PRT; 814 AA.
AC 065715;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 91.9 KD PROTEIN.
GN TSK18.220.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPRALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RA BEVAN M., VAN DER SCHUEREN J., CHUANG Y.-J., VOET M., ROBBEN J.,
RA VOLCKAERT G., BANCROFT I., MEWES H.W., MAYER K.F.X., SCHUELLER C.;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA EU ARABIDOPSIS SEQUENCING PROJECT;
RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AL022580; E1287635; -.
KW HYPOTHETICAL PROTEIN.
SO SEQUENCE 814 AA; 91943 MW; A41B8987 CRC32;

Query Match
Best Local Similarity 3.5%; Score 108; DB 10; Length 814;
Matches 21; Conservative 26; Mismatches 32; Indels 5; Gaps 5;

Db 729 EHSKNVHPNRTTYVMAIGYARGVNTFASRLNEMREKGIYPD-SITTYKEFIYGIK 787
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 187 EKXIQIYVXKQIYTVSDAVKNPGDVFODTIYVEDLKRGISAEKRLVYISV-AYXKQ 245
| : : : : | : : : : | : : : : | : : : : | : : : : |
Db 788 GGV-LEAFKSGDE-EN-YAIIISG 808
| : : : : | : : : : | : : : : | : : : : | : : : : |
OY 246 VYKLETTYSXSXEVAFAEALIKG 269
| : : : : | : : : : | : : : : | : : : : | : : : : |

```

```

RESULT 15
ID 067578 PRELIMINARY; PRT; 426 AA.
AC 067578;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FLAGELLAR HOOK ASSOCIATED PROTEIN FLICK.
GN FLGK.
OS AQUIFEX AECOLICUS.
OC BACTERIA; AQUIFICALES; AQUIFICACEAE; AQUIFEX.
RN [1]
pe SEQUENCE FROM N.A.
RC STRAIN-VF5:
RX MEDLINE: 98196666.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
   aecolicus."
RL NATURE 392:353-358(1998).
RN [2]
RC SEQUENCE FROM N.A.
RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AE000749; G2983977; -.
RD FLAGELLA.
SQ SEQUENCE 426 AA; 49567 MW; 47417AE0 CRC32;

Query Match 3.4%; Score 106; DB 2; Length 426;
Best Local Similarity 23.7%; Pred. No. 4.47e-01;
Matches 22; Conservative 24; Mismatches 39; Indels 8; Gaps 7;

DB 1 MEGASEQUALEV-YKKM-IDVKNRNANAOEENTVAEEPVVQSDLYSGITFQEVRI 58
OY 164 KFGSPFEKXGNSLDIDFNSVHSGEKXIOIVXKQIYYTVSVDAVK-N-PGDVFQDTVTV 220
DB 59 QNFNFNTRNEK-LSYVSYLE-ERRDYLSKLES 89
OY 221 EDLKORGISAERPLVYISXVAIXROVYL-KLET 252

```

Search completed: Mon Aug 30 16:12:37 1999
 Job time : 81 secs.

(TM)

Distribution rights by Oxford Molecular Ltd

```
Run on:      Mon Aug 30 16:02:26 1999;  MasPar time 18.51 Seconds
```

Residuals (n) 148

Residues 61, 148, 195
33, 46, 83, 239
257 = Xaa

Sequence:

1 MANKAVNDEILAMNYDKKL.....TISIWGTTLYPQVEDKVEN 47

Scoring table: PAM 150

Gap 11

Searched: 170751 seqs, 21266608 residues

```
Post-processing: Minimum Match 0%
Listing first 45 summaries
```

Database:

a-geneseq35

1.pat1 2.pat2 3.pat3 4.pat4 5.pat5 6.pat6 7.pat7
8.pat8 9.pat9 10.pat10 11.pat11 12.pat12 13.pat13
14.pat14 15.pat15 16.pat16 17.pat17 18.pat18
19.pat19 20.pat20 21.pat21 22.pat22 23.pat23
24.pat24 25.pat25 26.pat26 27.pat27 28.pat28
29.pat29 30.pat30 31.pat31 32.pat32 33.pat33
34.pat34 35.pat35 36.pat36 37.pat37 38.pat38
39.pat39

Statistics: Mean 35.335; Variance 169.953; scale 0.208

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred.	No
1	3239	99.9	471	1	R05923	Immunogenic pneumolys	3.78e-27	
2	3063	94.5	471	1	R05924	Immunogenic pneumolys	2.07e-25	
3	1425	43.9	480	6	R33842	Streptolysin O varian	1.96e-11	
4	1403	43.4	480	6	R33841	Soluble streptolysin	7.62e-11	
5	1398	43.1	571	2	R10376	Streptolysin O deriva	3.17e-10	
6	1104	34.1	371	1	R06000	Bacteriophage lambda	2.51e-83	
7	367	11.3	113	33	R60952	Streptococcus pneumoniae	8.63e-30	
8	112	3.5	11	17	R88135	Streptococcus pneumoniae	6.06e-100	
9	114	3.5	1228	17	R77673	S-layer protein encoded	4.45e+00	
10	114	3.5	1228	28	W22862	Bacillus stearothermophilus	4.45e+00	
11	108	3.3	188	29	W55347	H. pylori ORF 01ce116	1.12e+01	
12	108	3.3	299	33	W60976	Streptococcus pneumoniae	1.12e+01	
13	108	3.3	418	29	W55558	H. pylori ORF 29ge303	1.12e+01	
14	99	3.1	527	33	W61278	Monamine oxidase A.	4.32e+01	
15	99	3.1	527	2	R20579	Human monomelic oxidase	4.32e+01	
16	101	3.1	667	22	W20753	H. pylori transporter	3.21e+01	

[illegible]

ID R3841 standard; Protein; 480 AA.
 AC R3841;
 DT 08-JUL-1993 (first entry)
 DE Soluble streptolysin O variant.
 KW Recombinant; haemolytic activity; immunodiagnostic activity; ss.
 OS Streptococcus pyogenes.
 PN MO9305156-9.
 PD 18-MAR-1993.
 PF 31-JUL-1992; 006398.
 PR 30-AUG-1991; US-752429.
 PA (BECI) BECKMAN INSTR INC.
 PI Adams CW, Mang EV;
 DR WPI: 93-100980/12.
 DR N-PSDB: Q38286.
 FI
 FI Deriv. of streptolysin O with haemolytic activity - used in
 FI immuno-diagnostic assays which rely on haemolytic activity of
 FI wild-type Streptolysin O
 PS Claim 7; Fig 2; 60pp; English.
 CC The sequence is that of a soluble deriv. of Streptolysin O
 CC with haemolytic properties which was obt'd. from Streptococcus pyogenes
 CC genomic library clone rSO.3. It is useful in immunodiagnostic assays
 CC which rely upon, e.g. the haemolytic properties of wild type
 CC streptolysin O. Recombinant Streptolysin O is obt'd. more cheaply
 CC than purified Streptolysin O obt'd. from Streptococcus pyogenes.
 QJ Sequence 480 AA;

ID	RESULT	5
AC	R10376: standard: Protein; 571 AA.	
DT	05-APR-1991 (first entry)	
DE	Streptolysin O derivative.	
KW	SLO.	
OS	Streptococcus pyogenes.	
PN	GB2233977-A.	
PD	23-JAN-1991.	
PF	04-JAN-1989; 000107.	
PR	04-JAN-1989; GB-000107.	
PA	(KEHO/) KEHOE M.	
PI	(PINK/) PINKNEY M.	
DR	PI: kehoe M, pinkney M.	
DR	WPI: 91-024598/04.	
DR	N-PSDB: Q10320.	
PT	Deriv. of thiol-activated protein streptolysin O - contg. no	
PT	cysteine amino acid but retaining cytolytic activity, used for	
PS	detecting antibodies in samples	
PS	Disclosure; Fig 2: 9pp: English.	
CC	The SLO derivative contains no cysteine residues, with cytolytic	
CC	activity giving resistance to inactivation by oxidation or	
CC	thiol groups. Abs raised to the SLO may be used in the detection and	
CC	diagnosis of Streptococcus pyogenes infection.	
SC	Sequence 571 AA;	
Query Match	43.1%; Score 1398; DB 2; Length 571;	
Best Local Similarity	41.8%; Pred. No. 3.17e-109;	
Matches	191; Conservative 104; Mismatches 161; Indels 1; Gaps 1;	
Db	109 indktylanynelevlaknetienfypkqevkkadkfivierkkknitpvdids 168	
Qy	6 VNDPILAMNDKKLLTHGSEIENRFXKEGNLPDEFVYXKKKSLSTNSDXVAT 65	
Db	169 vctdtyaaqlangkyfdenkpdavvtckrrpqklnhdipmgdkat-vevndpyanvt 227	
Qy	66 NDSRLYRGALLVDEFLKKNPTLLAVDRAPMTYSIDLPGLASSDSFLOVEDPSSSVRG 125	
Db	228 aidhlyngvhdnygsncltpartqteamyksqgleaalnyvnsklglgtigdfktsk 287	
Qy	126 AVNDLAKWHODYGVGNVPAKQYERKTIHSHMDQLKVFSGDFEKTGNSLDIDFNSVHS 185	
Db	288 gekkymiaaykqglfvtysaanjpmnadvfdksvtlfeklgrkgvnsaaprlfysnvaygt 347	
Qy	186 GEEQIOIIVXKQIYITVSVDMKKNPGDVFQDFVTVEEDLKQRGISERPLVYISXVAGRG 245	
Db	348 vfwkletsksndvaafsaalkgtdtkngkyssdlmssstlavllvggdaenkvvtk 407	
Qy	246 VYLKLETTSKS-EVNAALFEALIKKVAAPOTEEKQILNDTEVKAAILLGGDSSGARVVTG 305	
Db	408 dfvlnrvnkdaatsrknpayrlsytsvflknkilaagvnnrteyvetsteytsgklnl 467	
Qy	306 KQVMWEDLLIOESRFPADHPGLPISTYTSPLRDNVAVAFQNSTDVEYKATVANGDLL 365	
Db	468 shggaaayagellwelnhdgkgevitkrrydmnysktsfsvlpjlgansnrlima 527	
Qy	366 DHSGAIVAOYITTNMELSYDHOGKEVLTPEAPMDRNGODTLTAFTTSIPLKGNVNLVSKI 425	
Db	528 reatglawewrkylderdvklskelnhvnsgtislsp 564	
Qy	426 RECTGLAEWMRTVEKTDLPVKKRTISLWGTITLP 462	
RESULT	5	
ID	R06000 standard: protein; 371 AA.	
AC	R06000;	
DT	22-NOV-1990 (first entry)	
DE	Bacteriophage lambda PL promoter - streptolysin O (SLO) fusion	
DE	protein of plasmid pMK306	
KW	Streptolysin O; SLO; protease; bacteriophage lambda; ds.	
OS	Streptococcus pyogenes.	
PH	key	
PT	Location/Qualifiers	
	cleavage_site 33..34	

CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or expected
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from *H. pylori* by PCR
CC amplification for recombinant polypeptide production, e.g. in *E. coli*
CC hosts.
SO Sequence 188 AA;

Query Match 3.3%; Score 108; DB 29; Length 188;
Best Local Similarity 34.4%; Pred. No. 1.12e+01;
Matches 21; Conservative 20; Mismatches 17; Indels 3; Gaps 3;

Db 95 psalmrdsvelskkr-1naikdlfhnkaf-rq1qklnp1kal-veaqqdgetka 151
Qy 210 PGDFVQDTVVEDLKRGISAEPLVYSXVAYGROYLKLETSKSEVEAFALING 269
Db 152 1 152
Qy 270 V 270

RESULT 12
ID W60976 standard; protein; 299 AA.

AC W60976; 13-OCT-1998 (first entry)
DE Streptococcus pneumoniae encoded polypeptide.
KW coding region; ORF; open reading frame; antibacterial;
KW infection; prevention; meningitis.
OS Streptococcus pneumoniae.
PN W09819689-A1.
PD 14-MAR-1998; 019226
PF 27-OCT-1997; 019226
PR (SMK) SMITHKLINE BEECHAM CORP.
PA (SMK) SMITHKLINE BEECHAM PLC.
PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO,
PI Reid RH, Zarfos PN;
PI WPI: 98-285586/25.
DR N-PSDB; V37377.
PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
PT e.g. for identifying anti-bacterial(s) for treatment and prevention
of meningitis
PS Claim 11: Page 71-72; 130pp; English.
CC The sequence is that of the polypeptide encoded by a region isolated
CC from *S. pneumoniae* which shows homology to nicotinate-nucleotide pyro-
CC phosphoriylase. It, or agonists of it, may be useful as an anti-
CC bacterial for treatment or prevention of infection, specifically caused
CC by *S. pneumoniae* (particularly meningitis) but possibly also *Helicobacter*
CC pylori (ulcers and gastric cancer). It may be of particular
CC use before insertion of an in-dwelling device or any other
CC invasive procedure. The protein, or nucleic acid encoding
CC it, can also be used in vaccines to induce a cellular
CC and/or humoral immune response, or to screen for other
CC antibacterials. The DNA may also contain flanking sequences
CC that are potential sources of control elements for bacterial
CC gene expression. Detecting a sequence encoding the protein
CC can be used diagnostically, e.g. to detect a mutation for
CC serotyping or classifying infectious agents.
SO Sequence 299 AA;

Query Match 3.3%; Score 108; DB 33; Length 289;
Best Local Similarity 21.4%; Pred. No. 1.12e+01;
Matches 22; Conservative 35; Mismatches 40; Indels 6; Gaps 6;

Db 38 dystalafdhgqkvsifakeagvlag-1tvfgvftlidaevtfqnpqfkgdgrlts 96
Qy 14 NDKKKLTHGCESTENFXKGNGLDPEFVXERKRSLSTWTS-DIXVATNDSRLYP 72
Db 97 gdl-vl-eiig-svrslltcervalnfnqhsiglasmaayve 136
Qy 73 GALIVDETLKNNPTLLAVDRAPMTYSIDLPGAS-SDSFLQ 114

RESULT 13
ID W55528 standard; protein; 418 AA.

AC W55528; 02-JUL-1998 (first entry)
DE *H. pylori* ORF 299g30321_24336712-fl-5 cellular protein.
KW cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
OS *Helicobacter pylori*.
PN W09737044-A1.
PD 09-OCT-1997;
PF 27-MAR-1997; 005223
PR 06-DEC-1996; US-761318.
PR 29-MAR-1996; US-625811.
PR 25-APR-1996; US-758731.
PR 25-OCT-1996; US-736905.
PR 28-OCT-1996; US-738859.
PA (ASTR) ASTRA AB.
PI Alm RA, Smith D;
PI WPI: 97-503122/46.
DR N-PSDB; V24937.

PT *Helicobacter pylori* nucleic acid sequences and encoded
PT polypeptides) - useful in vaccines to treat or prevent *H. pylori*
PT infection and for diagnosis of *H. pylori* infection
PS Claims 14, 93; Page 730-731; 1145pp; English.
CC This sequence is a *H. pylori* cellular protein.
CC The protein may be used in a vaccine to prevent or treat *H. pylori*
CC infection or to identify *H. pylori* polypeptide binding compounds,
CC useful as potential *H. pylori* life cycle activators or inhibitors. The
CC DNA and probes derived from it may be used for the identification of
CC *H. pylori* in a sample and the diagnosis of *H. pylori* infection. Nucleic
CC acid sequences complementary to the DNA act as antisense sequences and
CC can be used to prevent the translation of *H. pylori* mRNA. Antibodies
CC against the protein can be used in immunoassays to evaluate the abundance
CC and distribution of *H. pylori*-specific antigens. The genomic sequence
CC of *H. pylori* (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were analysed
CC for ORF of at least 180 nucleotides, and the predicted coding regions
CC defined by computer evaluation. To identify likely *H. pylori* antigens for
CC vaccine development, the amino acid sequences predicted from various ORF
CC were analysed for significant homology to other known or expected
CC membrane proteins. Having identified and determined the sequences of
CC interest, particular regions can be isolated from *H. pylori* by PCR
CC amplification for recombinant polypeptide production, e.g. in *E. coli*
CC hosts.
SO Sequence 418 AA;

Query Match 3.3%; Score 108; DB 29; Length 418;
Best Local Similarity 34.4%; Pred. No. 1.12e+01;
Matches 21; Conservative 20; Mismatches 17; Indels 3; Gaps 3;

Db 325 psalmrdsvelskkr-1naikdlfhnkaf-rq1qklnp1kal-veaqqdgetka 381
Qy 210 PGDFVQDTVVEDLKRGISAEPLVYSXVAYGROYLKLETSKSEVEAFALING 269
Db 382 1 382
Qy 270 V 270

RESULT 14
ID W61278 standard; protein; 527 AA.

AC W61278; 29-SEP-1998 (first entry)
DE Monamine oxidase A.
KW Monamine oxidase A; abnormal behaviour; serotonin; dopamine;
KW noradrenaline; mental retardation; impulsive aggression.
OS Homo sapiens.
PN US5783680-A.
PD 21-JUL-1998.
PF 06-OCT-1993; 132168.
PR 06-OCT-1993; US-132168.
PA (GEHO) GEN HOSPITAL CORP.

PA (UKA-) UNIV STICHTING KATHOLIEKE.
 PI Breakerfield XO, Brunner HG;
 DR WPI: 98-427102/36.
 DR N-PSDB: V27798.
 PT Genetic diagnosis of impulsive aggression - uses mutant forms of
 PS monoamine oxidase enzyme
 PS Disclosure: Column 47-50; 30pp; English.
 CC A genetic defect linked to p11-p21 of the X chromosome, especially
 CC mutation C936T in exon 8 of the monoamine oxidase (MAO) gene is related to
 CC abnormal behaviour. MAO metabolises serotonin, dopamine and
 CC noradrenaline, neurotransmitters whose metabolism is known to affect
 CC behaviour. Disregulation of these transmitters leads to borderline
 CC mental retardation and abnormal behaviour (including impulsive
 CC aggression). Probes based on the mutated sequence can be used for the
 CC genetic diagnosis of abnormal behaviour.
 SQ Sequence 527 AA;

Query Match 3.1%; Score 99; DB 33; Length 527;
 Best Local Similarity 19.3%; Pred. No. 4.32e+01;
 Matches 29; Conservative 40; Mismatches 75; Indels 6; Gaps 5;

DB 1 mengkasiagmfdvvnvlggsgisaak-llteygsvlvleardvgrtyltineh 59
 QY 250 LETTSKSEVEAFEPALIKGVKAPOTEMKQILDTEVKAVILGDPSSGARVYTKYDM 309
 DB 60 v-dyvdvggagvptqnrlrlskelgietykvnseilvgykgtypfgafpvpwmp 118
 QY 310 VEDLIOEGSRFTADHPGLPLSYTSTFLRDNV-VATFQNSTDYETKVTAYRNGDLLDHS 368
 DB 119 layl-dynmlwrti--dmngkeiptdapwe 145
 QY 369 GAVVAQYITWNELSYDHQKEVLTTPKAMD 398

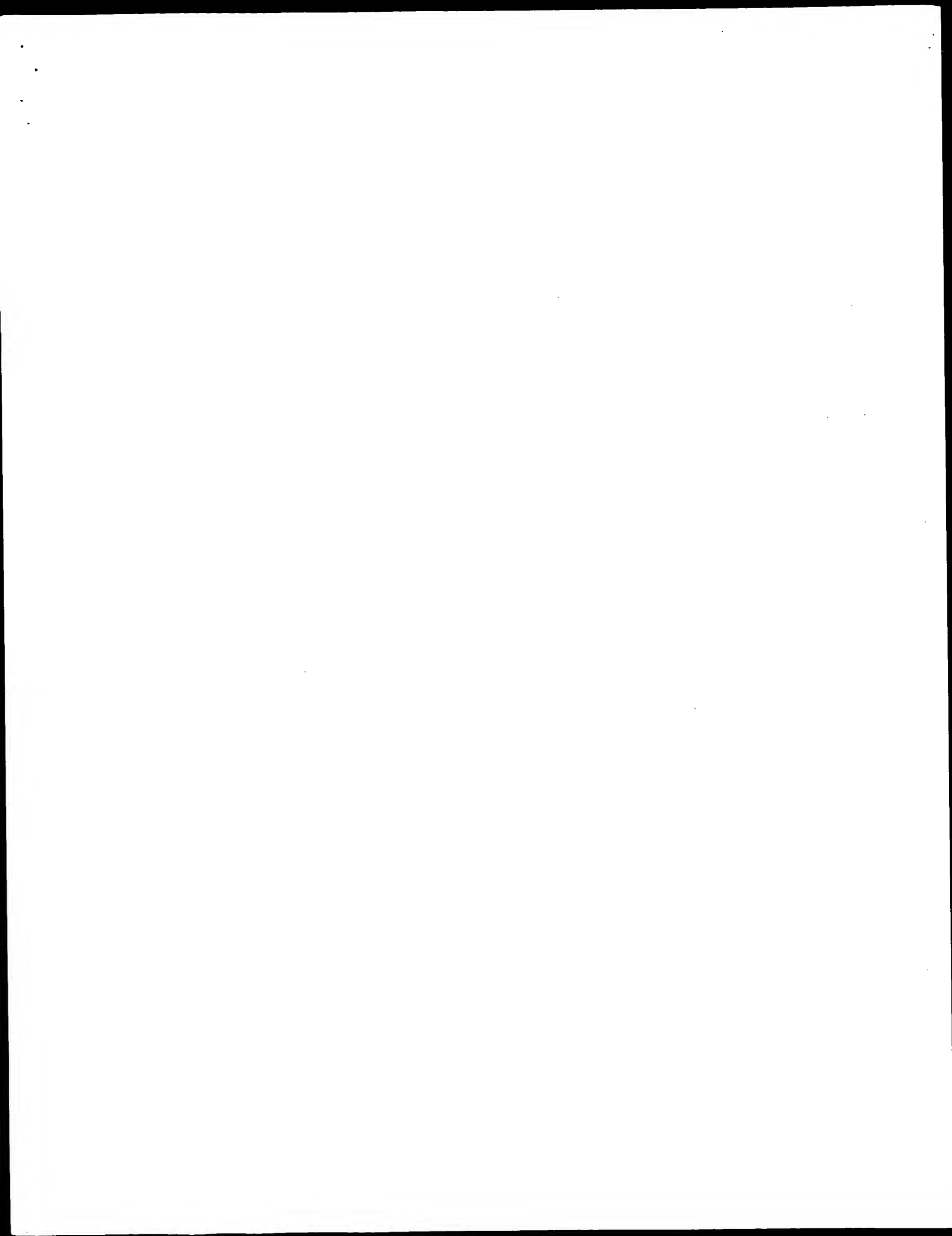
RESULT 15
 ID R05079 standard; protein: 527 AA.
 AC R05079:
 DT 10-MAR-1993 (revised)
 DT 11-JUL-1990 (first entry)
 DE Human monoamine oxidase type A gene product.
 KW Human monoamine oxidase type A; MOA-A; mental retardation;
 KW manic depression; psychotic disorders.
 OS Homo sapiens.
 PN MO9000195-A.
 PD 11-JAN-1990.
 PF 29-JUN-1989; 02901.
 PR 30-JUN-1988; US-213544.
 PA (BREA) Breakerfield X.
 PI Breakerfield X.
 DR WPI: 90-037130/05.
 DR N-PSDB: Q03217.
 PT Purified DNA encoding human monoamine oxidase type A -
 PT used to detect eg. MAO-A alterations, associated with manic
 PT depression, and expressed to form MAO inhibitors to treat psychotic
 PT disorders.
 PS Disclosure: Fig 2; 30pp; English.
 CC MAO-A catalyses oxidative deamination of dietary amines and
 CC neurotransmitters, eg. dopamine. The product may be used to develop
 CC therapeutic inhibitors, to metabolise dietary monoamines or to act as
 CC false transmitters.
 SQ Sequence 527 AA;

Query Match 3.1%; Score 99; DB 2; Length 527;
 Best Local Similarity 19.3%; Pred. No. 4.32e+01;
 Matches 29; Conservative 40; Mismatches 75; Indels 6; Gaps 5;

DB 1 mengkasiagmfdvvnvlggsgisaak-llteygsvlvleardvgrtyltineh 59
 QY 250 LETTSKSEVEAFEPALIKGVKAPOTEMKQILDTEVKAVILGDPSSGARVYTKYDM 309
 DB 60 v-dyvdvggagvptqnrlrlskelgietykvnseilvgykgtypfgafpvpwmp 118
 QY 310 VEDLIOEGSRFTADHPGLPLSYTSTFLRDNV-VATFQNSTDYETKVTAYRNGDLLDHS 368

DB 119 layl-dynmlwrti--dmngkeiptdapwe 145
 QY 369 GAVVAQYITWNELSYDHQKEVLTTPKAMD 398

Search completed: Mon Aug 30 16:04:20 1999
 Job time : 114 secs.



(TM)

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

```

Mon Aug 30 10:01:41 1999;
MasPar time 20.82 seconds
906 663 Miller

```

Tabular output not generated.

applies/sec

```

Title: >US-09-120-044-3-COPY
Description: (1-471) from devil2004.pep
Perfect Score: 3241
Sequence: 1 MANKAVNDFILAMNYDKKTL.....

```

.....IISLWGIILYPQVEDKVEN 47J

scoring table: PAM 150
Gap 11

122810 seqs, 40068593 residues

```

Post-processing: Minimum Match 0%
Listing first 45 summaries

```

Database: pir60

statistics: Mean 50.333; Variance 110.368; scale 0.456

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3239	99.9	471	2	A28568	perfringolysin - Strepto	0.00e+00
2	1599	49.3	500	2	B43577	perfringolysin O prec	1.85e-27
3	1542	47.6	529	2	A43505	listeriolysin O precu	1.62e-25
4	1535	47.4	529	2	S24231	listeriolysin precurs	3.97e-25
5	1525	47.1	528	2	S22341	lyanolysin precursor	2.58e-25
6	1476	45.5	530	2	S22340	seeligelolysin - Lis	7.35e-24
7	1468	45.3	501	2	A37858	alveololysin - Bacillus	2.51e-24
8	1415	43.7	571	2	A43507	streptolysin O precu	3.52e-23
9	1336	41.2	485	2	I19863	hemolysin - Bacillus	4.43e-22
10	464	14.3	112	2	S47298	sulfolysin - Streptococ	2.01e-57
11	124	6.6	96	2	S47297	sulfolysin - Streptococ	3.17e-15
12	124	6.6	527	2	S03974	amine oxidase (flavin)	1.66e-02
13	114	3.5	462	1	S00552	mitochondrial proteas	2.64e-01
14	112	3.5	748	1	S95327	hypothetical protein	4.51e-01
15	114	3.5	1228	2	I70468	surface layer protein	2.64e-01
16	113	3.5	1361	3	T03415	s-layer protein - Cam	3.46e-01
17	111	3.4	187	2	C71140	hypothetical protein	5.89e-01
18	107	3.3	218	2	F64461	hypothetical protein	1.67e+00
19	106	3.3	265	2	S00497	hypothetical protein	2.16e+00
20	108	3.3	365	2	F71665	hypothetical protein	1.29e+00
21	108	3.3	418	2	G71952	hypothetical protein	1.29e+00
22	106	3.3	501	2	A69415	conserved hypthetica	2.16e+00
23	106	3.3	511	2	S58322	nucleolar protein NOP	2.16e+00

	24	107	3.3	5762	2	A41819	proline-rich peptides	1.67e+00
	25	103	3.2	3003	2	S70118	probable membrane pro	4.62e+00
	26	103	3.2	364	2	UC1188	GTP cyclohydrolase II	4.62e+00
	27	105	3.2	418	2	E64555	conserved hypothetica	2.79e+00
	28	105	3.2	426	2	A70444	flagellar hook associ	2.79e+00
	29	104	3.2	505	2	S09638	flagellin - salmonel	3.59e+00
	30	105	3.2	514	1	DMEBRT	threonine dehydratase	2.79e+00
	31	101	3.1	259	2	B69113	cell division inhibi	7.60e+00
	32	102	3.1	277	2	S68595	phosphate-specific tr	5.93e+00
	33	100	3.1	279	2	S42125	hypothetical protein	9.71e+00
	34	101	3.1	280	2	S56569	hypothetical protein	9.71e+00
	35	101	3.1	352	2	B70876	probable aminotransf	7.60e+00
	36	100	3.1	447	2	S75210	glycine hydroxymethyl	9.71e+00
	37	102	3.1	466	2	E68999	L-amino acid oxidase	5.93e+00
	38	101	3.1	513	2	A65249	hypothetical 58.2 kd	9.71e+00
	39	101	3.1	722	2	S64392	hypothetical protein	7.60e+00
	40	102	3.1	726	2	C64236	protein V (fcrv) homo	5.93e+00
	41	102	3.1	1035	2	A64686	cation efflux system	5.93e+00
	42	101	3.1	1035	2	D47831	cation efflux system	7.60e+00
	43	101	3.1	1123	1	WMBE77	UL37 protein - human	7.60e+00
	44	100	3.1	1235	1	D32433	VSG expression site-a	9.71e+00
	45	100	3.1	1744	2	JH0720	tanabin - African cia	9.71e+00

ALIGNMENTS

RESULT	ENTRY	#type complete
1	A28568	pneumolysin - Streptococcus pneumoniae
TITLE	#formal_name	Streptococcus pneumoniae
ORGANISM	#sequence_revision	19-Nov-1988 #text_change
DATE		21-Aug-1998
ACCESSIONS	A28568; S12829	
REFERENCE	Walker, J.A.; Allen, R.L.; Falmagne, P.; Johnson, M.K.; Boulnois, G.J.	
#authors	Infect. Immun. (1987) 55:1184-1189	
#journal	Molecular cloning, characterization, and complete nucleotide sequence of the gene for pneumolysin, the sulfhydryl-activated toxin of Streptococcus pneumoniae.	
#title		
#cross-references	MUID:87193109	
#accession	A28568	
##molecule_type	DNA	
##residues	1-471	
REFERENCE	S12829	
#authors	Mitchell, T.J.; Mendez, F.; Paton, J.C.; Andrew, P.W.; Boulnois, G.J.	
#journal	Nucleic Acids Res. (1990) 18:4010	
#title	Comparison of pneumolysin genes and proteins from Streptococcus pneumoniae types 1 and 2.	
#cross-references	EMBL:X52474; NID:G47403; PID:G47404	
#accession	S12829	
##status	preliminary	
##molecule_type	DNA	
##residues	1-471	
##cross-references	EMBL:X52474; NID:G47403; PID:G47404	
CLASSIFICATION	#superfamily dipeptide transport protein	
FEATURE		
2-471	#product pneumolysin #status predicted #label MAT	
SUMMARY	#length 471 #molecular_weight 52899 #checksum 8274	
Query Match	99.9%; Score 3239; DB 2; Length 471;	
Best Local Similarity	98.1%; Pred. No. 0.00e+00;	
Matches	462; Conservative 1; Mismatches 8; Indels 0; Gaps 0;	

Db 1 MANAANDFLIANNVDETLLENNPTLAVDRAPMYSIDLPGASSDSFLOVEDPSN 60
OY 1 MANAAVDLFILANNVDETKLLTHQGESIEENFKESNOJLPDEFVAKERRRSISTNTSDI 60
Db 61 SVATAINDSRILYPALLVDETLLLENNPTLAVDRAPMYSIDLPGASSDSFLOVEDPSN 120
OY 61 XVTATNDSTRILYPALLVDETLLKKNPPTLAVDRAPMYSIDLPGASSDSFLOVEDPSN 120

Db 121 SSVRGAVNDLLAKKHODYGVYNNPARKQYKRTAHSMEOUKYKFGSDPEKTSNLDIDF 180
 QY 121 SSVRGAVNDLLAKKHODYGVYNNPARKQYKRTAHSMEOUKYKFGSDPEKTSNLDIDF 180
 Db 181 NSVHGEKOIOIVNFKOITYTSDAVAKNPGDFQDYETVEDLKORGISARPLVYISV 240
 QY 181 NSVHGEKOIOIVNFKOITYTSDAVAKNPGDFQDYETVEDLKORGISARPLVYISV 240
 Db 241 AYGRQVYTLKLETTSSKDEVEAFALIKGVAPOTEMKQILDNTYKAVYLGSDPSSGA 300
 QY 241 AYGRQVYTLKLETTSSKDEVEAFALIKGVAPOTEMKQILDNTYKAVYLGSDPSSGA 300
 Db 301 RVTGKVMVEDLLOEGSRFTADHPGLPISTTSLRNVAVATONSDYETVETATARN 360
 QY 301 RVTGKVMVEDLLOEGSRFTADHPGLPISTTSLRNVAVATONSDYETVETATARN 360
 Db 361 GDLILDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 420
 QY 361 GDLILDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 420
 Db 421 LSVKIRECTGLAMEWRVYKTDLPYKRTISITWGTITLYQVBDKEND 471
 QY 421 LSVKIRECTGLAMEWRVYKTDLPYKRTISITWGTITLYQVBDKEND 471

RESULT 2
 ENTRY #type complete
 TITLE perfringolysin O precursor - Clostridium perfringens
 ORGANISM #formal_name Clostridium perfringens
 DATE 03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change 13-Sep-1998

ACCESSIONS
 #accession B43577
 #molecule_type DNA
 #residues 1-500 #label SHI
 #cross-references GB:M01080; NID:g144885; PID:g144886
 #note Translation of the nucleotide sequence is not complete

REFERENCE
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3235-3240
 #title Nucleotide sequence of the gene for perfringolysin O (theta-toxin) from Clostridium perfringens: significant homology with the genes for streptolysin O and pneumolysin.

#cross-references MUID:89032623
 #accession A34951
 #molecule_type DNA
 #residues 1-125, 'EA', 129-500 #label TWE
 #cross-references GB:M6704; NID:g144883; PID:g144884

REFERENCE
 #authors Tweten, R.K.
 #journal Infect. Immun. (1988) 56:3228-3234
 #title Cloning and expression in Escherichia coli of the perfringolysin O (theta-toxin) gene from Clostridium perfringens and characterization of the gene product.

#cross-references MUID:85032622
 #accession A60922
 #molecule_type protein
 #residues 29-45 #label TW2
 #experimental_source ATCC 13124

GENETICS
 #gene pfo: pfoA

FEATURE
 1-28 #domain signal sequence #status predicted #label SIG
 29-500 #product perfringolysin O #status experimental #label MAT

SUMMARY
 #length 500 #molecular_weight 55799 #checksum 4821

Query Match 49.3%; Score 1599; DB 2; Length 500;
 Best Local Similarity 46.5%; Pred. No. 1,85e-270;
 Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;

Db 35 NOSIDSGJSSLSYNNNEVLASNGDKIESFVPEKKKAGNKRTIYVEROKSLTSPVDSI 94
 QY 35 NOSIDSGJSSLSYNNNEVLASNGDKIESFVPEKKKAGNKRTIYVEROKSLTSPVDSI 94
 Db 3 NKAVDFTILAMYYDKRLTLHOGESIENRFXKEGNQLPDEVEYERKKRSLSTSDIX 62
 QY 3 NKAVDFTILAMYYDKRLTLHOGESIENRFXKEGNQLPDEVEYERKKRSLSTSDIX 62
 Db 95 IDSVDRTYPPALADLADFAFENRPTILMYKRPININIDPLGKGENS-IKVDPTIYK 153
 QY 95 IDSVDRTYPPALADLADFAFENRPTILMYKRPININIDPLGKGENS-IKVDPTIYK 153
 Db 63 TATNDSRLTPGALLVYDELAKENPTLLAVRAPMTYSIDPLGLASSDSFQVDPSSNS 122
 QY 63 TATNDSRLTPGALLVYDELAKENPTLLAVRAPMTYSIDPLGLASSDSFQVDPSSNS 122
 Db 154 VSGAIDELVSKNNKYSSTHITPARTQYSESVMYKSGSISALANNAVLENSIGVDENA 213
 QY 154 VSGAIDELVSKNNKYSSTHITPARTQYSESVMYKSGSISALANNAVLENSIGVDENA 213
 Db 123 VRGAVNDLLAKKHODYGVYNNPARKQYKRTAHSMEOUKYKFGSDPEKTSNLDIDFNS 182
 QY 123 VRGAVNDLLAKKHODYGVYNNPARKQYKRTAHSMEOUKYKFGSDPEKTSNLDIDFNS 182
 Db 214 VANNEKWMILAYKQIEFYVSADLPKRNPSDLDEDSVFETNDLKOKSVNEADPLMVSNAV 273
 QY 214 VANNEKWMILAYKQIEFYVSADLPKRNPSDLDEDSVFETNDLKOKSVNEADPLMVSNAV 273
 Db 183 VHSGEKOIOIVNFKOITYTSDAVAKNPGDFQDYETVEDLKORGISARPLVYISXVAY 242
 QY 183 VHSGEKOIOIVNFKOITYTSDAVAKNPGDFQDYETVEDLKORGISARPLVYISXVAY 242
 Db 274 GRTIYVLETTSSSKDVOAFKALIKNDIKNSQYKDIYENSTFAYVVGDAQEHKVV 333
 QY 274 GRTIYVLETTSSSKDVOAFKALIKNDIKNSQYKDIYENSTFAYVVGDAQEHKVV 333
 Db 243 GRQVYTLKLETTSSKDEVEAFALIKGVAPOTEMKQILDNTYKAVYLGSDPSSGARV 302
 QY 243 GRQVYTLKLETTSSKDEVEAFALIKGVAPOTEMKQILDNTYKAVYLGSDPSSGARV 302
 Db 334 VTKDPEIRKIVKDNATFSTKNPAPISYTSVFLKDNVAAVHNKTDYIETSTYSGK 393
 QY 334 VTKDPEIRKIVKDNATFSTKNPAPISYTSVFLKDNVAAVHNKTDYIETSTYSGK 393
 Db 303 VTKGVMVEDLLOEGSRFTADHPGLPISTTSLRNVAVATONSDYETVETATARN 362
 QY 303 VTKGVMVEDLLOEGSRFTADHPGLPISTTSLRNVAVATONSDYETVETATARN 362
 Db 394 INDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 453
 QY 394 INDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 453
 Db 363 LLDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 422
 QY 363 LLDHSGAVAYOYITWDELSDYHOGKEVLTTPKAMRNGODLTAHFTTSIPKGVNRN 422
 Db 454 IKAECTGLAMEWRVYKTDLPYKRTISITWGTITLYQVBDKEND 493
 QY 454 IKAECTGLAMEWRVYKTDLPYKRTISITWGTITLYQVBDKEND 493

RESULT 3
 ENTRY #type complete
 TITLE listeriolysin O precursor - Listeria monocytogenes
 ORGANISM #formal_name Listeria monocytogenes
 DATE 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 17-Mar-1999

ACCESSIONS
 #accession A43505
 #molecule_type DNA
 #residues 1-529 #label MEN
 #cross-references GB:M24199; NID:g149652; PID:g149653
 #note this sequence is derived from a strongly hemolytic strain, serotype 1/2c

REFERENCE
 #authors Mengaud, J.; Vicente, M.F.; Chenevert, J.; Pereira, J.M.; Geoffroy, C.; Gicquel-Sauze, B.; Baquero, F.; Perez-Diaz, J.C.; Cossart, P.
 #journal Infect. Immun. (1988) 56:766-772
 #title Expression in Escherichia coli and sequence analysis of the listeriolysin O determinant of Listeria monocytogenes.

#cross-references MUID:88366684
 #accession S05306
 #molecule_type DNA
 #residues 1-529 #label DOM
 #cross-references EMBL:X15127; NID:g44106; PID:g44107
 #experimental_source strain EGD
 #note this sequence is derived from a weakly hemolytic strain, serotype 1/2a

REFERENCE
 A47606

```

#authors      Mengaud, J.; Chenevert, J.; Geoffroy, C.; Gallard, J.L.;
#journal      Infect. Immun. (1987) 55:3225-3227
#title        Identification of the structural gene encoding the
              SH-activated hemolysin of Listeria monocytogenes:
              listeriolysin O is homologous to streptolysin O and
              pneumolysin.
#cross-references MIMD:88057627
#accession    A47606
#status       preliminary
#molecule_type DNA
#residues     413-480 ##label ME2
#cross-references GB:M29171
#accession    S12400
#authors      Michel, E.; Reich, K.A.; Favier, R.; Berche, P.; Cossart, P.
#journal      Mol. Microbiol. (1990) 4:2167-2178
#title        Attenuated mutants of the intracellular bacterium Listeria
              monocytogenes obtained by single amino acid substitutions
              in listeriolysin O.
#cross-references MIMD:91211627
#accession    S12400
#molecule_type DNA
#residues     483-493 ##label MTC
#experimental_source strain LO28, serotype 1/2c

GENETICS
#gene          listA
#classification #superfamily dipeptide transport protein
#keywords       virulence factor
#feature        1-25
#summary        #domain signal sequence #status predicted #label SIG\
                #product listeriolysin O #status predicted #label MAT
                #length 529 #molecular-weight 58688 #checksum 719

Query Match      47.6%; Score 1542; DB 2; Length 529;
Best Local Similarity 42.9%; Pred. No. 1,628-259;
Matches 200; Conservative 117; Mismatches 148; Indels 1; Gaps 1;

Db 61 IDKTYGLDYNKNNVLYHGDAVTNPPRRGYKDGNEYIVKRRKKSINONNADIOYVNA 120
QY 6 VNDFTILAMNDKKLLTHGSEIENRFXKGNQLPDEFVYXERKRSLSNTSDIYAT 65
121 ISSLTYPGALVANSSELVENQDVLVPRKDSLTSLIDLPGMTNODKRIYVKNATKSNVNN 180
47 66 NDSRLYPGALLVYDETLXENNPFLAVDRAPMTYSIDLPGGLASSDSFLOVEDPSSSVRG 125
181 AVNTLVERNNEKXAQAYPNVSAKIDYDEMAVSESQLAKFGTAFAVNNSLNVNGAIS 240
QY 126 AVNDLAKMHOYDGOVN-NVPAKQYKETAHSMEDLAKYKFGSDPEKTSIDIDPNSVH 184
126 241 EGMKOEVISFOIYYNVNNEPTRPREFGKAVTKEOLALGVNAENPPAYISSVAYGR 300
QY 185 SGEKQIQIVXKQIYYTVSDAVKNPGDVFODTYVEDLKQKISAEPLVYISXVAYGR 244
185 301 QVYLKLTSTNSHSTKKAAPDAVSGKSVGDELNIINNSSEKAVIYGSAKDEVOIID 360
301 245 QVYLKLTSTNSKXEVAAFEALIKGVKAPQTEMKQILDNTEVKAVILGGDSSGARVYT 304
245 361 GNLGDLRLIKGATFNRETPGVPIAYTTNPLKDNLAIVKNNSEIETTSKAYTDGKIN 420
QY 305 GAYDMEDELIOESKFTADHPGLPISTTSFLRDVNAVAFQNSDYVEVETKATARNQDL 364
305 421 IDHSGVAAQFNISWDEINVDGEINVOHKMNSNNKSKLAHFTSSLYLPGNANINIVY 480
421 365 LDHSGAVYAAQYITVNNELSYDHQKREVLTPKAMDNGODLTAHFTTSLPLKGNVNNLSVK 424
365 481 AKECTGLAMEMWRTVIDDRNLPLVKNRNISIGTTLIYKYSNKNYDN 526
481 425 IRECTGLAMEMWRTVEKTDPLVYKRTISIGTTLIYQVEDKVEN 470

```

```

ORGANISM      12067)
#formal_name  Listeria monocytogenes
#strain       12067
#date         22-Nov-1993 #sequence_revision 10-Nov-1995 #text-change
              13-Sep-1998
ACCESSIONS
#authors      S24231
#journal      S24230
#title        Rasmussen, O.F.; Beck, T.; Olsen, J.E.; Dons, L.; Rossen, L.
              Infect. Immun. (1991) 59:3945-3951
              Listeria monocytogenes isolates can be classified into two
              major types according to the sequence of the listeriolysin
              gene.
#cross-references MIMD:92040062
#accession    S24231
#status       nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues     1-529 ##label RAS
#cross-references EMBL:X60035; NID:944110; PID:944112
#experimental_source strain 12067, serotype 4b
#note         the nucleotide sequence was submitted to the EMBL Data
              Library, June 1991

GENETICS
#gene          listA
#classification #superfamily dipeptide transport protein
#keywords       virulence factor
#feature        1-25
#summary        #domain signal sequence #status predicted #label SIG\
                #product listeriolysin #status predicted #label MAT
                #length 529 #molecular-weight 58687 #checksum 47

Query Match      47.4%; Score 1535; DB 2; Length 529;
Best Local Similarity 42.7%; Pred. No. 3,588-258;
Matches 199; Conservative 117; Mismatches 149; Indels 1; Gaps 1;

Db 61 IDKTYGLDYNKNNVLYHGDAVTNPPRRGYKDGNEYIVKRRKKSINONNADIOYVNA 120
QY 6 VNDFTILAMNDKKLLTHGSEIENRFXKGNQLPDEFVYXERKRSLSNTSDIYAT 65
121 ISSLTYPGALVANSSELVENQDVLVPRKDSLTSLIDLPGMTNODKRIYVKNATKSNVNN 180
QY 66 NDSRLYPGALLVYDETLXENNPFLAVDRAPMTYSIDLPGGLASSDSFLOVEDPSSSVRG 125
181 AVNTLVERNNEKXAQAYPNVSAKIDYDEMAVSESQLAKFGTAFAVNNSLNVNGAIS 240
QY 126 AVNDLAKMHOYDGOVN-NVPAKQYKETAHSMEDLAKYKFGSDPEKTSIDIDPNSVH 184
126 241 EGMKOEVISFOIYYNVNNEPTRPREFGKAVTKEOLALGVNAENPPAYISSVAYGR 300
QY 185 SGEKQIQIVXKQIYYTVSDAVKNPGDVFODTYVEDLKQKISAEPLVYISXVAYGR 244
185 301 QVYLKLTSTNSHSTKKAAPDAVSGKSVGDELNIINNSSEKAVIYGSAKDEVOIID 360
301 245 QVYLKLTSTNSKXEVAAFEALIKGVKAPQTEMKQILDNTEVKAVILGGDSSGARVYT 304
245 361 GNLGDLRLIKGATFNRETPGVPIAYTTNPLKDNLAIVKNNSEIETTSKAYTDGKIN 420
QY 305 GAYDMEDELIOESKFTADHPGLPISTTSFLRDVNAVAFQNSDYVEVETKATARNQDL 364
305 421 IDHSGVAAQFNISWDEINVDGEINVOHKMNSNNKSKLAHFTSSLYLPGNANINIVY 480
421 365 LDHSGAVYAAQYITVNNELSYDHQKREVLTPKAMDNGODLTAHFTTSLPLKGNVNNLSVK 424
365 481 AKECTGLAMEMWRTVIDDRNLPLVKNRNISIGTTLIYKYSNKNYDN 526
481 425 IRECTGLAMEMWRTVEKTDPLVYKRTISIGTTLIYQVEDKVEN 470

```

```

RESULT      4
ENTRY       S24231 #type complete
TITLE       listeriolysin precursor - Listeria monocytogenes (strain

```

```

RESULT      5
ENTRY       S22341 #type complete
TITLE       ivanolysin precursor - Listeria ivanovii
ORGANISM    #formal_name Listeria ivanovii
DATE        07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change
              21-Aug-1998

```

[illegible][illegible]

Query Match 45.3%; Score 1468; DB 2; Length 501;
 Best Local Similarity 41.9%; Pred. No. 2,51e-245;
 Matches 190; Conservative 121; Mismatches 141; Indels 1; Gaps 1;

DB 44 IAGLNANREVLAIQGDQSSSEVPKKEGINSNGKFLVERDKKSLTSPVDSITNSR 103
 10 IIAANNDRKKLLTHOGESIEENRFKREGNOLPDEFVYXERKKRSLSTNSTSDIXATNDSR 69
 DB 104 IYPGAIALANKDFADNOPSLVMAARKPLDISLDPGLKNENT-IVQNPNGTSSAIDQ 162
 70 LEPGALLVDELFXENNPFLAVDAPMTYSIDLPGLASDSFQLQEDPSNSVGAAND 129
 DB 163 LVSTGEKYSSTHTLPARLOYAESVYSONQISSALNNAKVLNGLTGLIDENAVNGEKK 222
 130 ELAKHMOHYGVANNVPARQYKETAHSMQELKAFGSDFEKTGSLDIDRNSHSGKQ 189
 DB 223 VWVAAYKOFYTVSAGLPNNPSLDLDDSYTFAELARKGVSNAPPLMYSNVAAGRTIYVK 282
 190 IOIVAKKQIYTVSDAVANKPGDVFQDITYVEDLKORGISAEPLVYISXVAYGROYLX 249
 DB 283 LETTSKNDVQTAFLKLLNPNFSIQASGOYKDIYENSSTFAVLGGDAQTHNOVTKDFNV 342
 250 LETTSKSEVEAFALLKGVAVAPQTEMKQILDNTYKAVILGGDPSSGARVVTGKAYDM 309
 DB 343 IQSVTKDNAQSSKNPAPVISTYSVLKDNSTIAAVHNTYETETTESKGIKLDHSG 402
 310 VEDLQESRFTADHPGLPISTYTSFLRDNVAVATFQNSTDYETKRVNVRNGDLDLHSG 369
 DB 403 AYVAQFEYVMEFSDADGOEIVTFKSDGWRBNSAFSTEIPLPNAKIRIFARECT 462
 370 AYVAQYITWMLSTDHQGEVLPKAWDRNGODLTAFTTISPLKGNVRLSVKIRECT 429
 DB 463 GLAEMWRTVYDEVNVPASDINVSIMGTTLYP 495
 430 GLAEMWRTVYEKTDLPVKKRTISIMGTTLYP 462

RESULT 8
 ENTRY A43507 #type complete
 TITLE streptolysin O precursor - Streptococcus pyogenes
 ORGANISM #formal name Streptococcus pyogenes
 DATE 28-Oct-1992 #sequence_revision 28-Oct-1992 #text_change 09-Sep-1997

ACCESSIONS A43507
 REFERENCE A43507
 #authors Kehoe, M.A.; Miller, L.; Walker, J.A.; Boulinois, G.J.
 #journal Infect. Immun. (1987) 55:3228-3232
 #title Nucleotide sequence of the streptolysin O (SLO) gene: structural homologs between SLO and other membrane-damaging, thiol-activated toxins.

#cross-references MUID:88057628
 #accession A43507
 #status preliminary
 #molecule_type DNA
 #residues 1-571 #label KEH
 #cross-references GB:M18638; NID:9153810; PID:9153811
 SUMMARY #length 571 #molecular_weight 63638 #checksum 73

Query Match 43.7%; Score 1415; DB 2; Length 571;
 Best Local Similarity 42.0%; Pred. No. 3.52e-235;
 Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

DB 109 INDKIYSLNNELEVLANKGETIENFVPEKGVKADKFIYTERKKKININTPVDISIDS 168
 6 VNDPILANNYDKKLLTHOGESIENRFKREGNOLPDEFVYXERKKRSLSTNSTSDIXAT 65
 DB 169 VTDSTYAAIALANKGETENKPDPAVVTKRNPKIHDLPFGMGKAT-VEVNDPTVANST 227
 66 NDSRLYFGALLVDETLXENNPFLAVDRAPMTYSIDLPGLASDSFQLQEDPSNSVAG 125
 DB 228 AIDLVMQMDHNTSGGVTLPARTQYTESMYSKQIQAALNANVSKLIDGLIGDFESISK 287

QY 126 AYNDLAKKHODYGOVANNVPARQYKETAHSMQELKAFGSDFEKTGSLDIDFNSVHS 185
 DB 288 GEKKVMIAAYKQIYTVSANLPNNPADVEDKSYTFEELORKGVSNAPPLMYSNVAAGRT 347
 186 GEKQIOIVNKKQIYTVSDAVANKPGDVFQDITYVEDLKORGISAEPLVYISXVAYGQ 245
 DB 348 VFYKLEISSKSNDEAFSAALGTDYKTKNGKSIDLNSFTAVVLGGDAAEHKNVTK 407
 246 VYKLEISSKSNDEAFSAALGTDYKTKNGKSIDLNSFTAVVLGGDAAEHKNVTK 407
 DB 408 DFDVIRNVIDNATFSSKKNPAPVISTYSVLKDNSTIAAVHNTYETETTESKGIKLDHSG 467
 306 KYDMVEDLQESRFTADHPGLPISTYTSFLRDNVAVATFQNSTDYETKRVNVRNGDLDL 365
 DB 468 SHQAVYAQYITWMLSTDHQGEVLPKAWDRNGODLTAFTTISPLKGNVRLSVKIRECT 527
 366 DHGAVYAQYITWMLSTDHQGEVLPKAWDRNGODLTAFTTISPLKGNVRLSVKIRECT 425
 DB 528 RECTGLAEMWRTVYDEVNVPASDINVSIMGTTLYP 495
 426 RECTGLAEMWRTVYEKTDLPVKKRTISIMGTTLYP 462

RESULT 9
 ENTRY 139863 #type fragment
 TITLE hemolysin - Bacillus cereus (fragment)
 ORGANISM #formal name Bacillus cereus
 DATE 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996

ACCESSIONS 139863
 REFERENCE 139863
 #authors Yotsudo, T.; Okumura, K.; Iwasaki, M.; Hara, A.; Kakitani, S.; Maehara, W.; Igarashi, H.; Hinuma, Y.
 #journal Infect. Immun. (1994) 62:4000-4004
 #title The gene encoding a new mitogenic factor in a Streptococcus pyogenes strain is a distributed only in group A streptococci.

#cross-references MUID:94341910
 #accession 139863
 #status preliminary; translated from GB/EMBL/DBJ
 #molecule_type DNA
 #residues 1-485 #label RES
 #cross-references GB:D21270; NID:9418066; PID:9600252
 SUMMARY #length 485 #checksum 3782

Query Match 41.2%; Score 1336; DB 2; Length 485;
 Best Local Similarity 39.8%; Pred. No. 4.43e-220;
 Matches 173; Conservative 117; Mismatches 144; Indels 1; Gaps 1;

DB 44 SIDTIGNLTYNNQEVLANVGKVESFVPKESINSNGKFFVYVYDRKHLQKHQSIIFRLD 103
 5 AYNDPILANNYDKKLLTHOGESIENRFKREGNOLPDEFVYXERKKRSLSTNSTSDIXAT 64
 DB 104 SVANRTYPGAVALANKAFADNOPSLVMAARKPLDISLDPGLKNENT-IVQNPNGTSSAIDQ 162
 65 TNDKRLYFGALLVDETLXENNPFLAVDRAPMTYSIDLPGLASDSFQLQEDPSNSVAG 124
 DB 163 GAVDDLVSTWNEKYSATHTLPARMQYTESMYSKQIQAALNANVSKLIDGLIGDFESISK 287
 125 GAVNDLAKKHODYGOVANNVPARQYKETAHSMQELKAFGSDFEKTGSLDIDFNSVH 184
 DB 222 NGEKVMIAAYKQIYTVSANLPNNPADVEDKSYTFEELORKGVSNAPPLMYSNVAAGRT 347
 185 SEKQIOIVNKKQIYTVSDAVANKPGDVFQDITYVEDLKORGISAEPLVYISXVAYGQ 245
 DB 283 IYVYKLEISSKSNDEAFSAALGTDYKTKNGKSIDLNSFTAVVLGGDAAEHKNVTK 407
 246 VYKLEISSKSNDEAFSAALGTDYKTKNGKSIDLNSFTAVVLGGDAAEHKNVTK 407
 DB 343 KDFNIRNVIDNATFSSKKNPAPVISTYSVLKDNSTIAAVHNTYETETTESKGIKLDHSG 467
 305 GRVMDVEDLQESRFTADHPGLPISTYTSFLRDNVAVATFQNSTDYETKRVNVRNGDLD 364

```

Db      403 LDHGAVAAQDVADWDSQFETPQNKELTFRKWTGSGSKDKTAHSTFYIPLPPNSKNIKIV 462
OY      365 LDHGAAYAAQYITVWNLSEIDHOGKEVLTPRKAMDNRQDULTAHFTTSIPILGNVRNLSVK 424
        463 ARECTGLAMEMWRTI 477
OY      425 IRECTGLAMEMWRTV 439

RESULT  10
ENTRY   #type complete
TITLE   suilysin - Streptococcus suis
ORGANISM #formal_name Streptococcus suis
DATE     06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
         09-Sep-1997

ACCESSIONS
REFERENCE
#authors Segers, R.
#submission submitted to the EMBL Data Library, September 1994
#accession #molecule_type DNA
          ##residues 1-112 ##label SRG
          ##cross-references EMBL:Z36908; NID:g535307; PID:g535308
SUMMARY  #length 112 #molecular_weight 12464 #checksum 9264

Query Match      14.3%; Score 464; DB 2; Length 112;
Best Local Similarity 50.9%; Pred. No. 2,01e-57;
Matches 57; Conservative 22; Mismatches 33; Indels 0; Gaps 0;

Db      1 PCVPPISTSTPYXXXPRXQXLNSNYEITITSTVHNSSMLTLXHSGAVAKYNINMGVSV 60
OY      325 PCLPISTYSTSPFLRDVVVFQNSTDYEFKKYAVARNGDLLDHSGAAYAQTIVWNLST 384
        61 NEAGEEVEXKAKMDKNGVNLTSHWSETIXIGNAXLNVNIQECTGLAMEWM 112
OY      385 DHQGEVLPKMKMDRNGODULTAHFTTSIPILGNVRNLSVKIRECTGLAMEWM 436

RESULT  11
ENTRY   #type complete
TITLE   suilysin - Streptococcus suis
ORGANISM #formal_name Streptococcus suis
DATE     06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
         09-Sep-1997

ACCESSIONS
REFERENCE
#authors Segers, R.
#submission submitted to the EMBL Data Library, September 1994
#accession #molecule_type DNA
          ##residues 1-96 ##label SEG
          ##cross-references EMBL:Z36907; NID:g535305; PID:g535306
SUMMARY  #length 96 #molecular_weight 10471 #checksum 6662

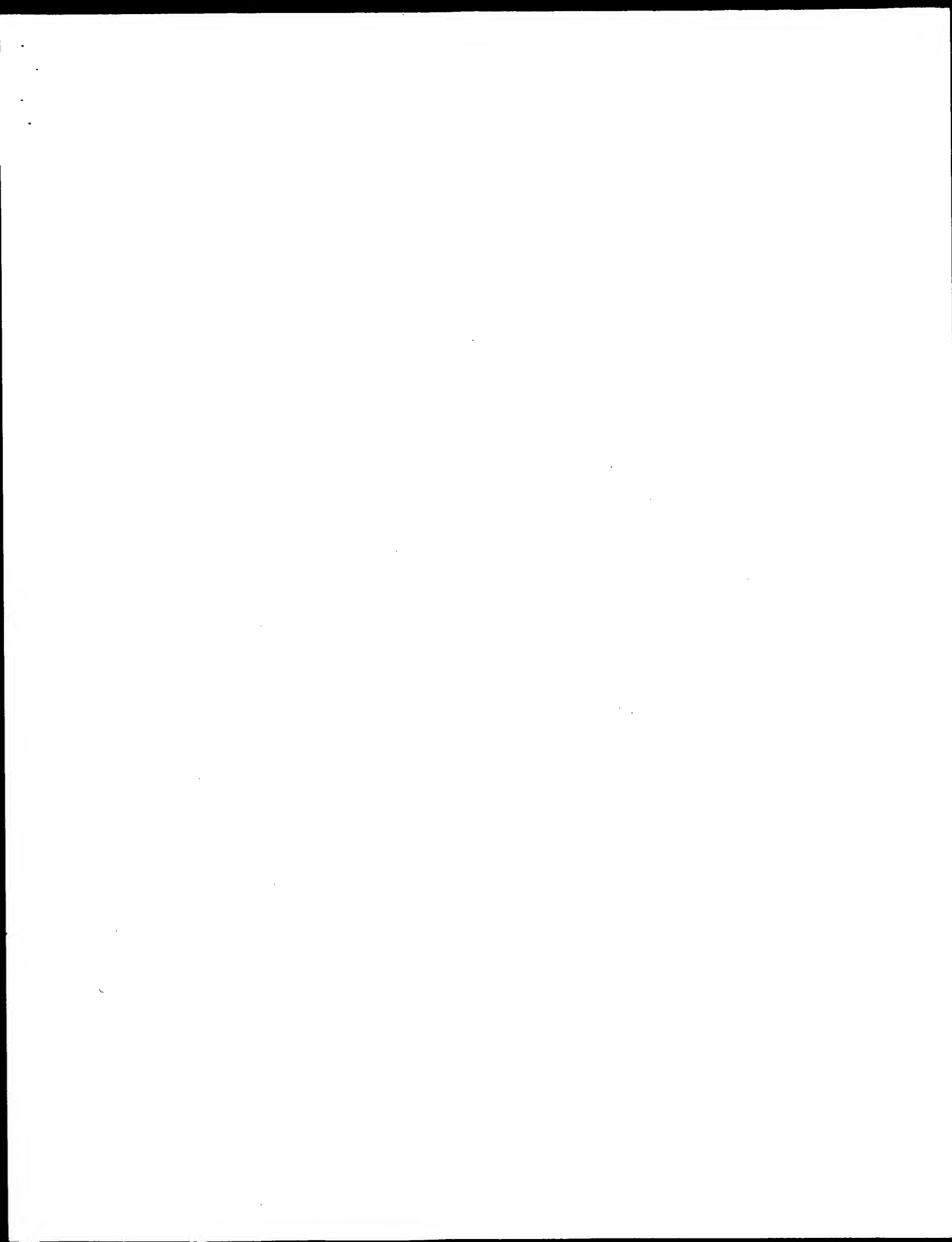
Query Match      6.6%; Score 214; DB 2; Length 96;
Best Local Similarity 33.7%; Pred. No. 3,17e-15;
Matches 30; Conservative 23; Mismatches 36; Indels 0; Gaps 0

Db      1 KODNOYFOSLYTGPOEILINEGEYIDNPATYGMLENGRVALLREKXTITDNSADITVI 60
OY      4 KAVNDPTLAMNNYKRKLTHQGESIEINRPXKKGNOQLPDEFVYXERKKRSLSSTNTSDIXVT 63
        61 XAKAANISPGALLRAKONLLDNNPITLI 89
OY      64 ATNDSRLYPGALLVDELTENNPITLAV 92

RESULT  12
ENTRY   #type complete
TITLE   amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine
ALTERNATE_NAMES monoamine oxidase type A

```

[illegible]



Matches 461; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Db 1 ANKAVNDFIAMVYDKKLLTHOGESIEENFKEGNQLPDEFVYERKRSLSTNTSDIS 60
 2 ANKAVNDFIAMVYDKKLLTHOGESIEENFKEGNQLPDEFVYERKRSLSTNTSDIX 61

Db 61 VTATNDSRLYPGALLVVDDELLENPTLLAVDRAPMYISIDLPGLASSDSFLOVEDPSSNS 120
 62 VTATNDSRLYPGALLVVDDELLENPTLLAVDRAPMYISIDLPGLASSDSFLOVEDPSSNS 121

Db 121 SVGAANDLLAKKHODYGOVNNVPAKQYKTAHSMEDOLKAFGSDPFKTSNIDIPDN 180
 122 SVGAANDLLAKKHODYGOVNNVPAKQYKTAHSMEDOLKAFGSDPFKTSNIDIPDN 181

Db 181 SVHSGEKQIOIVAFKOITTVSVDAVKNPGDVFQDTVVEDIKORGISAEPLVYISSVA 240
 182 SVHSGEKQIOIVAFKOITTVSVDAVKNPGDVFQDTVVEDIKORGISAEPLVYISSVA 241

Db 241 YGRQVYLKLETTSKSDEVEAFALIKGVVAPQTEMKOILNTEVKAVITLGGDESSGAR 300
 242 YGRQVYLKLETTSKSDEVEAFALIKGVVAPQTEMKOILNTEVKAVITLGGDESSGAR 301

Db 301 VVTKGVDMEDLIQESRFTADHPGLPISTYTSFLDNVATFONSTDVETKVTAYRNG 360
 302 VVTKGVDMEDLIQESRFTADHPGLPISTYTSFLDNVATFONSTDVETKVTAYRNG 361

Db 361 DILLDSGAVVAYITTWDELSTDHOGKVELLPKAMDNGODLJAFTTSLPKGNVRL 420
 362 DILLDSGAVVAYITTWDELSTDHOGKVELLPKAMDNGODLJAFTTSLPKGNVRL 421

Db 421 SVKIRCTGLAMEMWRTVEKTDLPVRRKRTISIMGTLLYPOVEKVENMD 470
 422 SVKIRCTGLAMEMWRTVEKTDLPVRRKRTISIMGTLLYPOVEKVENMD 471

QY 422 SVKIRCTGLAMEMWRTVEKTDLPVRRKRTISIMGTLLYPOVEKVENMD 471

RESULT 2
 ID TACY CLOPE STANDARD: PRT: 500 AA.
 AC P19995:
 DF 01-FEB-1991 (REL. 17, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DE 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DR PERRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).
 GN PFO OR PFOA OR PFOA.
 OS CLOSTRIDIUM PERFRINGENS.
 OC PLASMID PRT1B.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;
 OC CLOSTRIDIUM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91099951.
 RA SHIMIZU T., OKABE A., MINAMI J., HAYASHI H.;
 RT "An upstream regulatory sequence stimulates expression of the
 RL perringolysin O gene of Clostridium perfringens.";
 RT INFECT. IMMUN. 59:137-142(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-44.
 RX MEDLINE: 89032623.
 RA TWETEN R.K.;
 RT "Nucleotide sequence of the gene for perringolysin O (theta-toxin)
 RL from Clostridium perfringens: significant homology with the genes for
 RT streptolysin O and pneumolysin.";
 RL INFECT. IMMUN. 56:3235-3240(1988).
 RN [3]
 RP SEQUENCE OF 29-45 AND 305-312.
 RX STRAIN=PB6K;
 PY MEDLINE: 87076517.
 RA OHNO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.;
 RT "Cold-labile hemolysin produced by limited proteolysis of theta-toxin
 RL from Clostridium perfringens.";
 RL BIOCHEMISTRY 25:6048-6053(1986).
 RN [4]
 RP SEQUENCE OF 492-500 FROM N.A.

RC STRAIN=NCTC 8237;
 RX MEDLINE: 96123363.
 RA SHIMIZU T., KOBAYASHI T., BA-THIEN M., OHTANI K., HAYASHI H.;
 RT "Sequence analysis of flanking regions of the pfoA gene of
 RL Clostridium perfringens: beta-galactosidase gene (pbg) is located in
 RL the 3'-flanking region.";
 RL MICROBIOL. IMMUNOL. 39:677-686(1995).
 RN [5]
 RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
 RX MEDLINE: 88004463.
 RA IWAMOTO M., OHNO-IWASHITA Y., ANDO S.;
 RT "Role of the essential thiol group in the thiol-activated cytolysin
 RL from Clostridium perfringens.";
 RL EUR. J. BIOCHEM. 167:425-430(1987).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE: 97113505.
 RA FELL S.C., ROSSJOHN J., ROHDE K., TWETEN R.K., PARKER M.W.;
 RT "Crystallization and preliminary x-ray analysis of a thiol-activated
 RL cytolysin.";
 RL FEBS LETT. 397:290-292(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE: 97325744.
 RA ROSSJOHN J., FELL S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
 RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
 RL model of its membrane form.";
 RL CELL 89:685-692(1997).
 CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 CC EUKARYOTIC CELL MEMBRANES.
 CC -1- SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M81080; G144886; -;
 CC EMBL: M36704; G144884; -;
 CC EMBL: D49537; G1502275; -;
 CC PIR: B43577; B43577.
 CC DR PROSITE: P500481; THIOL_CYTOLYSINS; 1.
 CC DR PFAM: PF01289; ThiOL_cytolysin; 1.
 CC KW TOXIN; HEMOLYSIS; SIGNAL; PLASMID; 3D-STRUCTURE.
 CC FT SIGNAL 1 28
 CC FT CHAIN 29 500
 CC FT ACT SITE 459 459
 CC FT CONFLICT 30 30
 CC FT CONFLICT 34 34
 CC FT CONFLICT 126 128
 CC FT CONFLICT RKP -> I (IN REF. 3).
 CC FT CONFLICT RKP -> EA (IN REF. 2).
 CC SQ SEQUENCE 500 AA; 55799 MW; 17FC32BC CRC32;

Query Match 49.3%; Score 1599; DB 1; Length 500;
 Best Local Similarity 46.5%; Pred. No. 0.00e+00;
 Matches 214; Conservative 105; Mismatches 140; Indels 1; Gaps 1;

Db 35 NOSIDSGISLSYRNREVLASNGDKIESFVPRKGGKAGKFFIVEOROKSLTSPVDISI 94
 3 NKAIVNDFIAMVYDKKLLTHOGESIEENFKEGNQLPDEFVYERKRSLSTNTSDIX 62

Db 95 IDSVDNRYPGALQADKAFAVENPTLLVKKRPINIDLPGLGENS-IKVDPPYGR 153
 63 TATNDSRLYPGALLVVDDELLENPTLLAVDRAPMYISIDLPGLASSDSFLOVEDPSSNS 122

Db 154 VSGAIDELVSKNNEKSSHTLTPARTQYSESVMYKSKQSLANNAVAVENSLGVDPNA 213
 154 VSGAIDELVSKNNEKSSHTLTPARTQYSESVMYKSKQSLANNAVAVENSLGVDPNA 213

```

QY 123 VARGAVNDLLAKMHODYGOVNNVPARXOYEKITASHMEOLKVKFGSPDEKTSGLDIDENS 182
CC 214 VANNEKVMILATKOFJFYTSADLPKPNPSDLPDSDVTFNDLKQKGVSNAPPLMVSNAVY 273
CC 183 VHSGEKOIOLVNNKQIYTYVSDAVKNPQGVFODVTVEDLKORGISAERPLVYISXVAY 242
Db 274 GRTIYVLETTSSSKVOAFAKALIKNTDINKNOQKDIYENSSEFAVYLGDAQCHNRY 333
QY 243 GRQVYKLETTSSKXEVEAFELINGVKAPOLEMKQILDNTEKAVAILGDPSSGARY 302
Db 334 VTQDFEIRKIVYKDNATFSTKNPAPYISYTVFLKONSAAVANKDYIETSTESKCK 393
QY 303 VTGKVMVEDLIEGSRFTADHGLPISYTSFPLRDNVATFQNSDYETKVTAYARNCD 362
Db 394 INDHSGAVYAOFEVAMDEVSYDKEGNEVLTHKTWGNODKTAHSTVPLEANRNRY 453
QY 363 LLDHSGAVYAOYITWNELSYDHQKEVLTFRAMPDNGDHLAHTTSPILKGNVRLNS 422
Db 454 IKARECTGLAMEMWRDYISEYDPLNNINVSIMGTLLYP 493
QY 423 VKIRECTGLAMEMWRDYIEKTDLPVAKKRISTIMGTLLYP 462

RESULT 3
ID TACY_LISIV STANDARD; PRT; 529 AA.
AC P31831: 048747: 057096: 057206:
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JUN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
OS HLY OR LISA OR HLYA.
GN LISTERIA MONOCYTOGENES.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
LN LISTERIA.
RN [1]
RX SEQUENCE FROM N.A.
RC STRAIN-EGD / SEROTYPE 1/2A;
RX MEDLINE: 89366684.
RA DOMANN E., CHAKRABORTY T.;
RT "Nucleotide sequence of the listeriolsin gene from a Listeria
RT monocytogenes serotype 1/2a strain.";
RL NOCLEIC ACIDS RES. 17:6406-6406(1989).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE: 88153053.
RA MENDAUD J., VICENTE M.-F., CHENEVERT J., PEREIRA J.M., GEOFROY C.,
RA GICQUEL-SANZ B., BAQUERO F., PEREZ-DIAZ J.-C., COSSART P.;
RT "Expression in Escherichia coli and sequence analysis of the
RT listeriolsin O determinant of Listeria monocytogenes.";
RL INFECT. IMMUN. 56:766-772(1988).
RN [3]
RX SEQUENCE FROM N.A.
RC STRAIN-12067;
RX MEDLINE: 92040062.
RA RASMUSSEN O.F., BECK T., OLSEN J.E., DONS L., ROSSEN L.;
RT "Listeria monocytogenes isolates can be classified into two major
RT types according to the sequence of the listeriolsin gene.";
RL INFECT. IMMUN. 59:3945-3951(1991).
RN [4]
RX SEQUENCE FROM N.A.
RC STRAIN-F2365, F5782 (SEROTYPE 4B), AND F4233, F6789 (SEROTYPE 1/2B);
RA VINES A., SAMINATHAN B.;
SUBMITTED (JUL-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: SULHYDYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15127; G44107; -
DR EMBL; M24199; G149653; -
DR EMBL; X60035; G44112; -
DR EMBL; U25452; G887870; -
DR EMBL; U25443; G887016; -
DR EMBL; U25446; G887028; ALT_INIT.
DR EMBL; U25449; G887864; ALT_INIT.
DR PIR; S05306; S05306.
DR PIR; A43505; A43505.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSSP; P19995; 1PFO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1
FT CHAIN 25
FT ACT_SITE 26 529
FT ACT_SITE 484 484
FT VARIANT 35 35
FT VARIANT 39 39
FT VARIANT 438 438
FT VARIANT 523 523
SQ SEQUENCE 529 AA; 5868 MW; 269EA737 CRC32;
Query Match 47.6%; Score 1542; DB 1; Length 529;
Best Local Similarity 42.9%; Pred. No. 1,20e-300;
Matches 200; Conservative 117; Mismatches 148; Indels 1; Gaps 1;
Db 61 IDKTYOGLDYKNNVLYHGDAVTNVPKRGYKQGENIYVEKKKKSINONNADIQVYNA 120
QY 6 VDFETLANNYDKKLLTHQGESIENRFKKEGNOQLPDEFVYXERKRSLSTSTSDIXYAT 65
Db 121 ISLTYPGALVKAISELVEVDVLPVYKRDSTLSIDLPKMTDNKRIYVKNNAKSVNN 180
QY 66 NDSRLYPPALLVDETLLENPNFTLAVDAPRTYSIDLPGLASSDSFLQVEDPSNSSVRG 125
Db 181 AVNTLVERMNEKRYAQAPYNSAKIDYDDENAYSESQILAKFGTAFFKAVNNSLNAFSAIS 240
QY 126 AVNDLLAKMHODYGOVNNVPARXOYEKITASHMEOLKVKFGSPDEKTSGLDIDFNSVH 184
Db 241 ESKQGEVYISFKQIYVNVNNEPTRPSPRFGKAVTKRQLOALGVNAENPPAYISSVAYGR 300
QY 185 SEKQIOIIVNNKQIYTYVSDAVKNPQGVFODVTVEDLKORGISAERPLVYISXVAYGR 244
Db 301 QYTLKLTNSNSTKYKAFDAVSGSVSGDVELNTIINKSSFAVYIGSANKDEVQIID 360
QY 245 QYLLKLETTSSKXEVEAFELINGVKAPOLEMKQILDNTEKAVAILGDPSSGARYVT 304
Db 361 GNIQDLRLILKKGATFNETPGVPAYTNTLKNDELAVINKNESEIYETTSKATYDCKIN 420
QY 305 GKVDAVEDLIEGSRFTADHGLPISYTSFPLRDNVATFQNSDYETKVTAYARNDDL 364
Db 421 IDHSGVYAOFEVAMDEVNYPDEGNEIVQHKNNSENNKSLAHTSTSYLPGNARNINY 480
QY 365 IDHSGAVYAOYITWNELSYDHQKEVLTFRAMPDNGDHLAHTTSPILKGNVRLNSV 424
Db 481 AKRECTGLAMEMWRDYIDRNLPVKNRISTIMGTLLYPQVEKVEN 526
QY 425 IRECTGLAMEMWRDYIEKTDLPVAKKRISTIMGTLLYPQVEKVEN 470

RESULT 4
ID TACY_LISIV STANDARD; PRT; 528 AA.
AC P31831.
DT 01-JUL-1993 (REL. 26, CREATED)
DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE IVANOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).

```



```

QY      365 LDHSGAVYAOYITWMELSYDHOGKEVLTPKAMDNRGODLTAHFTTISIPAKGVNRLSYK 424
DB      482 AREDTGFEMWMTYVDDRLNPLVKNKNSVINGTTLTPRRSNVND 527
QY      425 IRECTGLAMWMTYVEKTDLPVKKRTISIMGTTLPQVEDAVEN 470

RESULT 6
ID TACY PAEAL STANDARD; PRT; 501 AA.
AC P23564;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ALVEOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ALV.
OS PARENTHACILLUS ALVEI (BACILLUS ALVEI).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
RN PARENTHACILLUS.
RP SEQUENCE FROM N.A.
RA GLOFFROY C., MENGAUD J., ALOUF J.E., COSSART P.;
RX MEDLINE; 91072294.
RT "Alveolysin, the thiol-activated toxin of Bacillus alvei, is homologous to listeriolysin O, pertussin toxin O, pneumolysin, and streptolysin O and contains a single cysteine."
RL J. BACTERIOL. 172:7301-7305(1990).
RN [2]
RP SEQUENCE OF 33-60.
RA ALOUF J.E., GEOFFROY C., GILLES A.M., FALMAGNE P.;
RT (IN) RAPPOUJI R., ALOUF J.E., FALMAGNE P. (EDS.);
RL BACTERIAL PROTEIN TOXINS, PP.49-50, GUSTAV FISHER VERLAG, STUTTGART (1990).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62709; G142473; -.
DR PIR; A37858; A37858.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSP; P19995; IPRO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 32
FT CHAIN 33 501
FT ACT_SITE 461 461 ALVEOLYSIN.
FT CONFLECT 56 56 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 501 AA; 35268 MW; 8DF2C94A CRC32;

Query Match 45.3%; Score 1468; DB 1; Length 501;
Best Local Similarity 41.9%; Pred. No. 3.32e-284;
Matches 190; Conservative 121; Mismatches 141; Indels 1; Gaps 1;

```

```

QY      130 LIAKHWDYGVNNVPARQYKAITAHSMEOIAKVFSGDEKTSNIDIDFNSVHSGEKQ 189
DB      223 VVVAAYKQIEYTVSAGLPNNPSDLFDDSVFAELARGVSEAPPLWNVAVAGRTIYK 282
QY      190 IQIVNKKQIYTVSDAVKAPGVDFQDVIYVEDLKQGISAEPLVITSXAAGROYLYK 249
DB      283 LETTSKNDVOTAEKLLNNPSIOASGGQKDIYENSSFTAVLGGDQATHNOVTKDFNV 342
QY      250 LETTSKSEVEAEAPALLKGVKVAPOREMKOIDLNTIEVKAVILGGDSSGARYVTGVADM 309
QY      310 VEDLIEGSRFTADHPGIPISYTSFTLSDNVVAFONSTQVETKVAAYRNGDILLDHSG 369
DB      403 AVVAQFEVYVMEFSDADGCEIVTRKSMWGWRBSAHFETIPLPPNKNIRIFARECT 462
QY      370 AVVAQYITWMELSYDHOGKEVLTPKAMDNRGODLTAHFTTISIPAKGVNRLSYK 429
DB      463 GLAMWMTYVEKTDLPVKKRTISIMGTTLP 495
QY      430 GLAMWMTYVEKTDLPVKKRTISIMGTTLP 462

RESULT 7
ID TACY STRPY STANDARD; PRT; 571 AA.
AC P21131;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
RN STREPTOCOCCUS.
RP SEQUENCE FROM N.A.
RA MEDLINE; 88057628.
RX KEHOE M.A., MILLER L., WALKER J.A., BOUINIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural homologies between SLO and other membrane-damaging, thiol-activated toxins."
RL INFECT. IMMUN. 55:3228-3232(1987).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M18638; G13811; -.
DR PIR; A43507; A43507.
DR PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSP; P19995; IPRO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMAID.
FT SIGNAL 1 33
FT CHAIN 34 571
FT ACT_SITE 530 530 STREPTOLYSIN O.
FT CONFLECT 571 571 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA; 63638 MW; 33124E54 CRC32;

Query Match 43.7%; Score 1415; DB 1; Length 571;
Best Local Similarity 42.0%; Pred. No. 1.91e-272;
Matches 192; Conservative 104; Mismatches 160; Indels 1; Gaps 1;

```

```

QY 6 VNDPFLAAMYDKKLLTHOGESIEENRFEKQOLPDEEVAXERKRSLSTSTSDIXVAT 65
DB 169 VMDRTYPAALQIANKGFTENKPDVAVVTRKPNOKIHIDLPKMGDKAT-VEVNDPTVANST 227
QY 66 NDSRLYPGALLVDEFTLENNPTLLAVRAPMTYSIDLPGIASSDSFLQVDEPSSVGR 125
DB 228 AIDNLVNMHNDYSGGNTLPARTQYTESMYSKSQIEALVANSKIDLTGIDFKSISK 287
QY 126 AVNDLLAKWHODYGOVNNVPARQYKELTASHMEQLKVFESDEKTNISIDIDFNSVHS 185
DB 288 GEKKVIAAYKOIETVSANLPNNPADYEDKSVTEKQORRGVNEAPPLFVSNAVAGRT 347
QY 166 GEKOQIYVXKOIYTVSDAKNKGDFODTVYEDLKQGISAEPLVYSXVANGRO 245
DB 348 VFKLETSSKSDVDEAPSAALKGTDTVKTNGKYSIDLENSFTAVVLGGDAEHNKVVTK 407
QY 246 VYKLETSSKXEVAAFEALIKGVKAPQTEWKQIIDLNTVEKAVILGGDPSGARVVTG 305
DB 408 DEDVIRNVIKDNATSRKNPAVPISTYSVFLKNNKIAGVNNRTEVETSTSTYSGKINTL 467
QY 306 KYDWEDEDIQESRFTADHPGLPISTYSFLRDVNVAFQNSTDYETVKVAYANGDLL 365
DB 468 SHOGAYVAQYELIMDEINVDKGEVITRRDNNWNSKTSPESTVIPLGANSNIRIMA 527
QY 366 DHSGAYVAQYITWELSYDHQGEVILPKAMDNRNGODLTHFTTSIPLGAVNANLSVKI 425
DB 528 RECTGLAEMWNRKVYDEEDVKLSKEINVINISGTLSP 564
QY 426 RECTGLAEMWNRKVYKTDLPVLRKRRTISIGTLLP 462

RESULT 8
ID TACY STREQ STANDARD; PRT; 574 AA.
AC Q54114;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SIO.
OS STREPTOCOCCUS EOUISIMILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SIMD-1;
RX MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D16824; G498301;
CC PROSITE; PS00481; THIOI_CYTOLYSINS; 1.
CC PFM; PF01289; Thiol_cytolysin; 1.
CC HSP; P19995; 1PFO.
CC TOXIN; HEMOLYSIN; CYTOLYSIS; SIGNAL.
CC SIGNAL 1 36 BY SIMILARITY.
FT CHAIN 37 574 STREPTOLYSIN O.

```

```

FT ACT_SITE 533 533 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 574 AA; 63991 MW; 9EBF31B4 CRC32;
Query Match 43.7%; Score 1416; DB 1; Length 574;
Best Local Similarity 41.8%; Pred. No. 1,14e-272;
Matches 191; Conservative 106; Mismatches 159; Indels 1; Gaps 1;
DB 112 INDKYISLNYNLEVLAKGFTIENFVREGVKKADKRTVIERKKNTPTPDISIDS 171
QY 6 VNDPFLAAMYDKKLLTHOGESIEENRFEKQOLPDEEVAXERKRSLSTSTSDIXVAT 65
DB 172 VMDRTYPAALQIANKGFTENKPDVAVVTRKPNOKIHIDLPKMGDKAT-VEVNDPTVANST 230
QY 66 NDSRLYPGALLVDEFTLENNPTLLAVRAPMTYSIDLPGIASSDSFLQVDEPSSVGR 125
DB 231 AIDNLVNMHNDYSGGNTLPARTQYTESMYSKSQIEALVANSKIDLTGIDFKSISK 290
QY 126 AVNDLLAKWHODYGOVNNVPARQYKELTASHMEQLKVFESDEKTNISIDIDFNSVHS 185
DB 291 GEKKVIAAYKOIETVSANLPNNPADYEDKSVTEKQORRGVNEAPPLFVSNAVAGRT 350
QY 186 GEKOQIYVXKOIYTVSDAKNKGDFODTVYEDLKQGISAEPLVYSXVANGRO 245
DB 351 VFKLETSSKSDVDEAPSAALKGTDTVKTNGKYSIDLENSFTAVVLGGDAEHNKVVTK 410
QY 246 VYKLETSSKXEVAAFEALIKGVKAPQTEWKQIIDLNTVEKAVILGGDPSGARVVTG 305
DB 411 DEDVIRNVIKDNATSRKNPAVPISTYSVFLKNNKIAGVNNRTEVETSTSTYSGKINTL 470
QY 306 KYDWEDEDIQESRFTADHPGLPISTYSFLRDVNVAFQNSTDYETVKVAYANGDLL 365
DB 471 SHOGAYVAQYELIMDEINVDKGEVITRRDNNWNSKTSPESTVIPLGANSNIRIMA 530
QY 366 DHSGAYVAQYITWELSYDHQGEVILPKAMDNRNGODLTHFTTSIPLGAVNANLSVKI 425
DB 531 RECTGLAEMWNRKVYDEEDVKLSKEINVINISGTLSP 567
QY 426 RECTGLAEMWNRKVYKTDLPVLRKRRTISIGTLLP 462

RESULT 9
ID TACY STREQ STANDARD; PRT; 574 AA.
AC Q53957;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SIO.
OS STREPTOCOCCUS CANIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SIMD-1;
RX MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----

```


Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Mon Aug 30 15:59:45 1999;  MasPar time 29.02 Seconds
Tabular output not generated.  885.876 Million cell updates/sec

```

```

File: >US-09-120-044-3-COPY
Description: (1-471) From devil20044 .pep
Perfect Score: 3241
Sequence: 1 MANAAVDFILAMNYDKKTL.....TISNGTLLPQVEDRYKEND 471

```

Scoring table: PAM 150

Searched: 179066 seqs, 54579741 residues

```
Post-processing: Minimum Match 0%
Listing first 45 summaries
```

Database: **sptrembl19**

1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 49.380; Variance 93.387; scale 0.5229

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Query No.	Score	Match Length	DB	ID	Description	Pred. No.
1	1779	54.9	497	2	055996	0.00e+00
2	1774	54.7	497	2	085102	0.00e+00
3	1302	40.2	534	2	031241	1.55e-24
4	124	3.8	50	2	048772	1.79e-03
5	118	3.6	50	2	048773	1.27e-02
6	117	3.6	450	2	068318	1.75e-02
7	112	3.5	748	2	012098	8.43e-02
8	113	3.5	1361	2	087083	6.18e-02
9	113	3.5	1361	2	030524	6.18e-02
10	111	3.4	187	1	058072	1.15e-01
11	111	3.4	814	10	065715	1.15e-01
12	107	3.3	218	1	058691	3.89e-01
13	106	3.3	285	3	012044	5.25e-01
14	106	3.3	501	1	028947	5.25e-01
15	107	3.3	710	1	085906	3.89e-01
16	107	3.3	5027	11	063460	3.89e-01
17	103	3.2	303	3	005541	1.27e+00
18	105	3.2	426	2	065758	7.07e-01
19	103	3.2	860	5	009647	1.27e+00
20	104	3.2	867	2	006185	9.51e-01
					SULFAXIN.	
					HEMOLISIN.	
					LYSTRIOLYSIN O (FRAGM	
					LYSTRIOLYSIN O (FRAGM	
					LISTERIOLYSIN O (FRAGM	
					HYPTHOETICAL 49.7 KD P	
					CHROMOSOME XII READING	
					S-LAYER PROTEIN.	
					S-LAYER PROTEIN.	
					187NA LONG HYPTHOETICA	
					HYPTHOETICAL 91.9 KD P	
					HYPTHOETICAL PROTEIN M	
					CHROMOSOME XV READING	
					CHROMOSOME XY READING	
					CONSERVED HYPTHOETICAL	
					TRANSPONSED PROTEIN B.	
					PROLINE-RICH PROTEIN B.	
					CHROMOSOME IV COSMID 9	
					FLAGELLAR HOOK ASSOCIA	
					HYPTHOETICAL 74.1 KD P	
					CLPB CHAPERONE HOMOLOG	

ALIGNMENTS

[illegible]

RN [2]

Best Local Similarity 18.8%; Pred. No. 8.43e-02;

OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
 OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
 OC CAPRIFALES; BRASSICACEAE; ARABIDOPSIS.

RA BEVAN M., VAN DER SCHUEREN J., CHUANG Y.-J., VOET M., ROBBEN J.,
 RA VOLCKERT G., BANCROFT I., MEWES H.W., MAYER K.F.X., SCHUELLER C.,
 RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]

RP SEQUENCE FROM N.A.

RA EU ARABIDOPSIS SEQUENCING PROJECT;

RL SUBMITTED (APR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; AL022580; E1287635; -

KW HYPOTHETICAL PROTEIN.

SO SEQUENCE 814 AA; 91943 MW; A418987 CRC32;

Query Match 3.4%; Score 111; DB 10; Length 814;
 Best Local Similarity 23.8%; Pred. No. 1.15e-01;

Matches 20; Conservative 29; Mismatches 30; Indels 5; Gaps 5;

Db 729 EMBL/GENBANK/DBJ DATA BANKS; 787
 QY 187 EMBL/GENBANK/DBJ DATA BANKS; 245

Db 788 GGV-LEA-FKGS-D-ENYALIEG 808
 QY 245 VYLKLETTSKSXEVAEAFELING 269

RESULT 12
 ID 058691 PRELIMINARY; PRT; 218 AA.
 AC 058691;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN M01295.
 GN M01295.
 OS METHANOCOCCUS JANNASCHII.
 OC ARCHAEA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE;
 OC METHANOCOCCUS.
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE; 96337999.
 RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
 RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
 RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.T.,
 RA OVERBEK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODER A.,
 RA SCOTT J.L., GEORGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
 RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
 RA COTTON M.D., ROBERTS C.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
 RA KLENK H.-P., FRASER C.M., SMITH H.O., MOESE C.R., VENTER J.C.,
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii";
 RL SCIENCE 273:1058-1073(1996).
 CC -1- SIMILARITY: STRONG, TO A.FULGIDUS AF1550 AND
 CC M.THERMOPHILICUM MTH576.
 CC EMBL; 067570; G1591933; -

DR HYPOTHETICAL PROTEIN.
 KW HYPOTHETICAL PROTEIN.
 SO SEQUENCE 218 AA; 24666 MW; CFC99A99 CRC32;

Query Match 3.3%; Score 107; DB 1; Length 218;
 Best Local Similarity 26.3%; Pred. No. 3.89e-01;

Matches 21; Conservative 17; Mismatches 40; Indels 2; Gaps 2;

Db 68 IGVGLTHIPKPSIPKAKKADLGAIEIVYH-GEVVEVEVEKTYTYSISDDIAT 126
 QY 153 ITHASHEQLVAKGSDPEKSGNSLDIDFNSVHSGEKQIOLVNAKQIYTVSDAVKPNPD 212

Db 127 PGFIDETAELENKENDIVE 146
 QY 213 V-FODTVTVEDLKRGISAE 231

RESULT 13
 ID 012044 PRELIMINARY; PRT; 265 AA.
 AC 012044;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE CHROMOSOME XV READING FRAME ORF YOR220W.
 OS SACHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCYCETES; SACHAROMYCETALES;
 OC SACHAROMYCETACEAE; SACHAROMYCES.
 RN [1]

RP SEQUENCE FROM N.A.
 RA BOYER J., FAIRHEAD C., GAILLON L., GALLISSON F., MICHAUX G.,
 RA THIERRY A., DUJON B.,
 RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]

RP SEQUENCE FROM N.A.

RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; X92441; G1050772; -

KW HYPOTHETICAL PROTEIN.

SO SEQUENCE 265 AA; 29255 MW; 2D5DD78E CRC32;

Query Match 3.3%; Score 106; DB 3; Length 265;
 Best Local Similarity 26.7%; Pred. No. 5.25e-01;

Matches 26; Conservative 26; Mismatches 44; Indels 7; Gaps 7;

Db 85 PKKLEVTSLNNQ-HPR-SRSTDVAISLODNNLALDEHKNKPLSTNPGYGVDS 142
 QY 12 ANV-YDKKKLTHQGESIENFKKGNLPDE-IVYAE-RKKRSISTSDIXVATINDS 68

Db 143 SUNKGSSLPKSSLESPYMKLSTDSKPSYQEPKLSRSS 187
 QY 69 RLY-GALLVYDELKXNNPTL-LAVDRAPMTYSIDLFGIASSDS 111

RESULT 14
 ID 028847 PRELIMINARY; PRT; 501 AA.
 AC 028847;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE CONSERVED HYPOTHETICAL PROTEIN.
 GN AF1322.
 OS ARCHAEoglobus fulgidus.
 OC ARCHAEA; EURYARCHAEOTA; ARCHAEoglobales; ARCHAEoglobaceae;
 OC ARCHAEoglobus.
 RN [1]

RP SEQUENCE FROM N.A.
 RA MEDLINE; 98049343.
 RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE O., NELSON K.E.,
 RA KETCHUM K.A., DODSON R.J., GWINN M., HICKY E.K., PETERSON J.D.,
 RA RICHARDSON D.L., KERLAVAGE A.R., GRAHAM D.E., KIRKNESS E.F.,
 RA FLEISCHMANN R.D., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
 RA PETERSON S., REICH C.T., MCNEIL L.K., BADGER J.H., GLODER A., ZHOU L.,
 RA OVERBEK R., GOCAYNE J.D., WEIDMAN J.F., MCDONALD L., UTTERBACK T.,
 RA COTTON M.D., SPRIGGS T., ARTIACH P., KAINE B.P., SYRES S.M.,
 RA SADOW P.W., D'ANDREA K.P., BOWMAN C., FUJII C., GARLAND S.A.,
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., MOESE C.R.,
 RA VENTER J.C.,
 RA "The complete genome sequence of the hyperthermophilic,
 RT sulphate-reducing archaeon Archaeoglobus fulgidus";
 RL NATURE 390:364-370(1997).
 DR EMBL; AE001012; G2649254; -

DR TIGR; AF1322; -

KW HYPOTHETICAL PROTEIN.


```

Db 181 nsyhsgeqigqivnfkgiyytvsdvavknpgdvfgdvtvedlkgqrgisaeprlvyissv 240
OY 181 NSVHSGEQIOIVNFKQIYTVSDAVKNPGDVFGDVTVEDLKQRGISAEPLVYISSV 240
Db 241 aysrgvylkletsksdeveaafaelikgvavapotekqkildntevkavilggdpsga 300
OY 241 AYSGRVYLKLETSSKDSDEVAEAFELIKGVAVAPOTEKQKILDNTEVKAVILGGDPSGA 300
Db 301 rvtgkvdmvedllqegsrftadhpjlpisyltsfirdnvaatqnsdtyvektayrn 360
OY 301 RVYTGKVMVEDLLQEGSRFTADHPGLPISTYTSFLRDNVVAATQNSDYEKVTAYRN 360
Db 361 gdlldhsgayvaoyiitwnelstyhockevltpkardnrcodtlafttsiplkgnvrn 420
OY 361 GDLLDHSGAYVAOYIITWNELSYDHOCKEVLTPKARNRCODTLAFTTSTIPLKGNVRN 420
Db 421 lsvklrectglawewrtvyekeidplvrkrttsiwtglttlypvedkvend 471
OY 421 LSVKRECTGLAWEMWRTVYEKIDPLVRKRTISIKGTTLTPYVEDKVEN 471

```

```

RESULT 2
ID R05924 standard; protein; 471 AA.
AC R05924;
DT 23-NOV-1990 (first entry)
DE Immunogenic pneumolysin variant.
KW pneumolysin; vaccine; pneumonia; meningitis; bacteraemia; ds.
OS Streptococcus pneumoniae.
FH Key Location/Qualifiers
FT misc-difference 428
FT /label=G, A, S
FT /label=Q, D
FT W09006951-A.
FN 28-JUN-1990.
PD 15-DEC-1989; AU0539.
PR 16-DEC-1988; AU-001989.
PA (PATO/) PATON J C.
PI PATON JC, HANSMAN DJ, BOULNOIS GJ, ANDREW PM, MITCHELL TJ,
PI WALKER JA;
DR WPI; 90-224494/29.
PT New non-toxic, immunogenic mutants of pneumolysin - useful in
PT protective vaccines against Streptococcus pneumoniae, and DNA
PT sequences encoding them.
PS Disclosure, P; English.
CC Vaccines are non-toxic and antigenic to wild type pneumolysin,
CC making them useful in vaccination against pneumonia and associated
CC meningitis, bacteraemia etc.
SQ Sequence 471 AA;

```

```

Query Match 94.6%, Score 3121; DB 1; Length 471;
Best Local Similarity 96.4%; Pred. No. 2.05e-259;
Matches 454; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```

```

Db 1 mankavdflamnydkkllthbgseierfikegnqlpdeveierkrrsiststsd 60
OY 1 MANKAVDFLAMNYDKKKLLTHBGSEIERFIKEGNQLPDEVEIERKRRSISTSTSDI 60
Db 61 svratndsrlypgallvvdetllennptllavdrapmtysidlpqlassdsfiqvedpsn 120
OY 61 SVRATNDSRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPGLASSDSFIQVDPSPN 120
Db 121 ssrvgaavndllakwqdgvgvnnparmyekktahameqkjkvfgsdfektgnsldidf 180
OY 121 SSVRGAAVNDLLAKWQDVGCVNNPARMOYEKKTASHAMEQLKVFSGSDEKTSNDIDF 180
Db 181 nsyhsgeqigqivnfkgiyytvsdvavknpgdvfgdvtvedlkgqrgisaeprlvyissv 240
OY 181 NSVHSGEQIOIVNFKQIYTVSDAVKNPGDVFGDVTVEDLKQRGISAEPLVYISSV 240
Db 241 aysrgvylkletsksdeveaafaelikgvavapotekqkildntevkavilggdpsga 300
OY 241 AYSGRVYLKLETSSKDSDEVAEAFELIKGVAVAPOTEKQKILDNTEVKAVILGGDPSGA 300

```

```

OY 241 AYSGRVYLKLETSSKDSDEVAEAFELIKGVAVAPOTEKQKILDNTEVKAVILGGDPSGA 300
Db 301 rvtgkvdmvedllqegsrftadhpjlpisyltsfirdnvaatqnsdtyvektayrn 360
OY 301 RVYTGKVMVEDLLQEGSRFTADHPGLPISTYTSFLRDNVVAATQNSDYEKVTAYRN 360
Db 361 gdlldhsgayvaoyiitwnelstyhockevltpkardnrcodtlafttsiplkgnvrn 420
OY 361 GDLLDHSGAYVAOYIITWNELSYDHOCKEVLTPKARNRCODTLAFTTSTIPLKGNVRN 420
Db 421 lsvklrectglawefrtvyekeidplvrkrttsiwtglttlypvedkvend 471
OY 421 LSVKRECTGLAWEMWRTVYEKIDPLVRKRTISIKGTTLTPYVEDKVEN 471

```

```

RESULT 3
ID R33842 standard; protein; 480 AA.
AC R33842;
DT 15-JUL-1993 (first entry)
DE Streptolysin O variant mSLO.3/6.
KW SLO; soluble; haemolytic activity; wild type; anti-SLO antibodies;
KW Streptococcus pyogenes; ASO.
OS Streptococcus pyogenes.
FH W09305155-A.
FN 18-MAR-1993;
PD 03-AUG-1992; U06380.
PR 30-AUG-1991; US-752428.
PA (BECT) BECKMAN INSTR INC.
PI Adams CW;
DR WPI; 93-100979/12.
DR N-FSDB; Q38287.
DT Streptolysin O variants produced by recombinant DNA technology -
DT having no haemolytic activity and recognised by wild-type
PT anti-streptolysin O antibodies, useful in diagnosis of
PT Streptococcus pyogenes infection
PS Claim 17, Fig 2, 80pp; English.
CC The sequence represents a soluble variant of Streptolysin O (SLO)
CC having no haemolytic activity. The variant SLO can specifically
CC recognise wild-type anti-SLO antibodies (ASO) but has no haemolytic
CC activity. It can be used for the determination of previous and/or
CC current infection by Streptococcus pyogenes. It enables ASO assays
CC to become commercially viable and avoids the risk of handling haemo-
CC lytically active S. pyogenes wild-type SLO.
SQ Sequence 480 AA;

```

```

Query Match 43.8%, Score 1444; DB 6; Length 480;
Best Local Similarity 42.8%; Pred. No. 7.18e-112;
Matches 196; Conservative 108; Mismatches 151; Indels 3; Gaps 3;

```

```

Db 18 indkiyslmynelevlakngeltien-fvpkegykackdifiavetkknknttpvdistsld 76
OY 18 INDKIYSLMYNELEVLAKNGELTIEN-FVPEKEGYKACKDIFIIVETKKNKNTTPVDISTSLD 76
Db 6 vndflamnydkkllthbgseierfikegnqlpdeveierkrrsiststsd 64
OY 6 VNDFLAMNYDKKKLLTHBGSEIERFIKEGNQLPDEVEIERKRRSISTSTSDI 64
Db 77 svdrtpaalqlnkftenkpdavvtknpgkibidpnmqdkat-vevndptyanvs 135
OY 77 SVDRTPAALQLNKFTENKPDVAVTKNPGKIBIDPQMDKAT-VEVNDPTYANVS 135
Db 65 tndsrlypgallvvdetllennptllavdrapmtysidlpqlassdsfiqvedpsn 124
OY 65 TNSDRLYPGALLVVDETLLENNPTLLAVDRAPMTYSIDLPGLASSDSFIQVDPSPN 124
Db 136 taidnlvqwhdnysgncltpartqytksmwysksgiaelnvnsklldgtldfksis 195
OY 136 TAINDLVQWHDNYSGNCLTPARTQYTKSMWYSKSGIAELNVNSKLLDGTLDFKSIS 195
Db 125 gavyndllakwqdgvgvnnparmyekktahameqkjkvfgsdfektgnsldidf 184
OY 125 GAVYNDLLAKWQDVGCVNNPARMOYEKKTASHAMEQLKVFSGSDEKTSNDIDF 184
Db 196 kgekvmaiaaykqifytvanljpnpadvfdksvtikelqjrkqvsneapplifsvnaaygr 255
OY 196 KGEKVMAIAAYKQIFYTVANLJPNPADVFDKSVTEKELQJRKQVSNAPPLIFSVNAYGR 255
Db 185 sgekoioivnfkgiyytvsdvavknpgdvfgdvtvedlkgqrgisaeprlvyissv 244
OY 185 SGEKOIOIVNFKQIYTVSDAVKNPGDVFGDVTVEDLKQRGISAEPLVYISSV 244
Db 256 tvfklletsksdeveaafaelikgvavapotekqkildntevkavilggdpsga 315
OY 256 TVFKLETSSKDSDEVAEAFELIKGVAVAPOTEKQKILDNTEVKAVILGGDPSGA 315
Db 245 ovytklettsksdeveaafaelikgvavapotekqkildntevkavilggdpsga 304
OY 245 OYTKLETSSKDSDEVAEAFELIKGVAVAPOTEKQKILDNTEVKAVILGGDPSGA 304
Db 316 kdfavirnvikdnatferknpayisyltsvflknkjavnnrtevetstetysgkin 375
OY 316 KDFAVIRNVIKDNATFERKNPAYISYLSVFLKNKJAVNNRTEVETSTETYSGKIN 375
Db 305 gkvmdvmedllqegsrftadhpjlpisyltsfirdnvaatqnsdtyvektayrn 420
OY 305 GKVMDEVEDLLQEGSRFTADHPGLPISTYTSFLRDNVVAATQNSDYEKVTAYRN 420

```


[illegible]

Pt	Nucleic acid encoding signal peptide of Bacillus stearothermophilus
Pt	S-layer protein - which has a lysine content of at least 10 per cent.
PS	Disclosure; Page 6-7; 12pp; German.
CC	The claimed signal peptide encoding sequence (see tag e of T08695)
CC	is pref. operably linked at the 3' end to a protein encoding
CC	sequence. The protein is pref. the S-layer protein (see tag f of
CC	T08695). At the 5' end, the signal peptide encoding sequence is
CC	pref. linked to an expression control sequence, pref. the sequence
CC	given in T08696.
SQ	Sequence 1228 AA;
Query Match	3.3%; Score 110; DB 17; Length 1228;
Best Local Similarity	24.7%; Pred. No. 9,46e+00;
Matches	41; Conservative 43; Mismatches 72; Indels 10; Gaps 10.
Dd	200 ydliv-amkar-ey-qdavkagnldkaavdginqlylpkvtdafktelt-evakkald 255 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy	200 YTVSVDVAKNNGDVDFODTIVEDLKKRGISIREPLVTYISSAVYGRQYYLKLETTSSK-DE 258
Dd	256 deaaltpkvesv-aaintqmkavelatavpyngt-lkqlgsaaamedvnnvtviyrkyvdg 313 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy	259 VEAAFEALIKGVKVAPOTEMKOI-LDITEXKAYVLIGDPSSGAAVWYGKDWMEDLLIOEG 317
Dd	314 nlpalntadavslstgkltlvydstpfenniteyykvvvkgjkdng 359 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy	318 S-RFLADHFGLPISITYTSFLDNVNVAIFONSTDY-VETKYTAIRANG 361

Query Match	3.3%	Score 110;	DB 25;	Length 1228;
Best Local Similarity	24.7%	Pred. No. 9.46e+00;		
Matches	41;	Conservative	43;	Mismatches 72; Indels 10; Gaps 10

Db 200 ydtlv-amkar-ev-qdaukaenldkakaavdinglylpkvdafkelt-evakkalda 255
 QY 200 YTVSVDAVKNPGGVFDVTVDLKGKRGISAERPLVYISSVAYGRQVYLKLETTSKS-DE 258
 Db 256 deaalpkrvesv-aintgpkaveltavpugt-1klqtsaanedtyvnyvrykxdg 313
 QY 259 VEAFAEALLKGVAVAGQTEMKQI-LDNTVEKAVILAGDSSGARVYTGKVDKVEDLIOEG 317
 Db 314 nlpfalntadsvlsldgkltvdaaspfennteykvvkgikdng 359
 QY 318 S-RTFADHFGPLSYTTSFLRDNVATFONSTDY-VEIKVATYRNG 361

RESULT 12
 ID W55347 standard; Protein; 188 AA.
 AC W55347;
 DT 17-JUN-1998 (first entry)
 DE H. pylori ORF 01cel1618orf11 protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 DR WPI: 97-503122/46.
 DR N-PSDB; V24937.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claim 14; Page 571; 1145pp; English.

CC This sequence is a H. pylori protein of unspecified function.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 188 AA;

Query Match 3.2%; Score 104; DB 29; Length 188;
 Best Local Similarity 34.4%; Pred. No. 2.32e+01;
 Matches 21; Conservative 21; Mismatches 16; Indels 3; Gaps 3;

Db 95 pslmrdsvsledskrr-inaikdlifhknkf-rqlqlkintpika-1veagdgdfka 151
 QY 210 PGDVPDVTVDLKGKRGISAERPLVYISSVAYGRQVYLKLETTSKSDEVAEFAALIKG 269
 Db 152 1 152
 QY 270 V 270

RESULT 13
 ID W5528 standard; Protein; 418 AA.
 AC W5528;
 DT 02-JUL-1998 (first entry)
 DE H. pylori ORF 299630321_24336712-f1_5 cellular protein.
 KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.
 OS Helicobacter pylori.
 PN W09737044-A1.
 PD 09-OCT-1997.
 PF 27-MAR-1997; U05223.
 PR 06-DEC-1996; US-761318.
 PR 29-MAR-1996; US-625811.
 PR 02-APR-1996; US-758731.
 PR 25-OCT-1996; US-736905.
 PR 28-OCT-1996; US-738859.
 PA (ASTR) ASTRA AB.
 PI Alm RA, Smith D.
 DR WPI: 97-503122/46.
 DR N-PSDB; V24937.
 PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection
 PS Claims 14,93; Page 730-731; 1145pp; English.

CC This sequence is a H. pylori cellular protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors. The
 CC DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance
 CC and distribution of H. pylori-specific antigens. The genomic sequence of
 CC H. pylori (ATCC 55679) was determined from overlapping contigs generated
 CC by mechanically shearing the bacterial DNA. The sequences were analysed
 CC for ORF of at least 180 nucleotides, and the predicted coding regions
 CC defined by computer evaluation. To identify likely H. pylori antigens for
 CC vaccine development, the amino acid sequences predicted from various ORF
 CC were analysed for significant homology to other known or exported
 CC membrane proteins. Having identified and determined the sequences of
 CC interest, particular regions can be isolated from H. pylori by PCR
 CC amplification for recombinant polypeptide production, e.g. in E. coli
 CC hosts.
 SQ Sequence 418 AA;

Query Match 3.2%; Score 104; DB 29; Length 418;
 Best Local Similarity 34.4%; Pred. No. 2.32e+01;
 Matches 21; Conservative 21; Mismatches 16; Indels 3; Gaps 3;

Db 325 pslmrdsvsledskrr-inaikdlifhknkf-rqlqlkintpika-1veagdgdfka 381
 QY 210 PGDVPDVTVDLKGKRGISAERPLVYISSVAYGRQVYLKLETTSKSDEVAEFAALIKG 269

Db 382 1 382
 QY 270 V 270

RESULT 14
 ID R80623 standard; Protein; 210 AA.
 AC R80623;
 DT 12-MAR-1996 (first entry)
 DE Coho salmon growth hormone.
 KW Growth hormone; coho salmon growth hormone; cSH; transgenic; abalone;
 KW actin; ACT5-C; insulin-like growth factor; cytochalasin B;
 KW polyploid organism.
 OS Oncorhynchus keta.
 FH Key
 FT peptide
 FT 1..22 Location/Qualifiers
 FT /note= "signal peptide"
 FT 23..210

 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Run on: Mon Aug 30 16:07:06 1999; MasPar time 20.87 Seconds
 Tabular output not generated. 904.473 Million cell updates/sec

Title: >US-09-120-044-3
 Description: (1-471) from US09120044.pep
 Perfect Score: 3299
 Sequence: 1 MANKAVNDFILAMNYDKKL.....TISMGITLYPOVEDKEND 471

Scoring table: PAM 150
 Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p1r60
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 50.559; Variance 111.373; scale 0.454

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3297	99.9	471	2	A28568	pneumolysin - Strepto	0.00e+00
2	1618	49.0	500	2	B43577	perfringolysin O prec	1.09e-772
3	1558	47.2	529	2	A43505	listeriolysin O precu	3.22e-261
4	1551	47.0	529	2	S24231	listeriolysin precurs	7.01e-261
5	1541	46.7	538	2	S22341	ivanolysin precursor	5.70e-250
6	1502	45.5	530	2	S22340	seeligeriolysin-Lis	1.55e-250
7	1492	45.2	501	2	A37858	alveolysin - Bacillus	1.28e-248
8	1436	43.5	571	2	A43507	streptolysin O precu	6.09e-238
9	1358	41.2	485	2	I39863	hemolysin - Bacillus	4.28e-223
10	464	14.1	112	2	S47298	streptolysin - Streptococ	4.77e-57
11	225	6.8	96	2	S47297	suilysin - Streptococ	8.99e-17
12	124	3.8	527	2	S03974	amine oxidase (flavin	1.92e-02
13	118	3.6	1361	3	T03415	S-layer protein - Cam	1.03e-01
14	117	3.5	265	2	S60947	hypothetical protein	1.35e-01
15	112	3.4	511	2	S58322	hypothetical protein NOP	5.16e-01
16	110	3.3	365	2	F71655	hypothetical protein	8.72e-01
17	109	3.3	744	2	A70385	DNA gyrase A subunit	1.13e+00
18	110	3.3	872	2	S73785	surfactant protein A	8.72e-01
19	110	3.3	1228	2	I40468	DNA polymerase III al	8.72e-01
20	106	3.2	187	2	S71140	hypothetical protein	2.44e+00
21	104	3.2	303	2	S70118	hypothetical protein	2.44e+00
22	104	3.2	418	2	G71952	probable membrane pro	4.04e+00
23	106	3.2	462	1	S00552	hypothetical protein	2.44e+00
						mitochondrial process	

24	106	3.2	501	2	A69415	conserved hypothetical	2.44e+00
25	106	3.2	1029	2	T02576	hypothetical protein	2.44e+00
26	102	3.1	277	2	S68595	phosphate-specific tr	6.65e+00
27	101	3.1	352	2	B70876	probable aminotransfe	8.50e+00
28	102	3.1	418	2	E64555	conserved hypothetical	6.65e+00
29	101	3.1	427	2	S75210	glycine hydroxymethyl	8.50e+00
30	102	3.1	446	2	E69899	L-amino acid oxidase	6.65e+00
31	103	3.1	512	2	S70644	annexin VII - African	5.19e+00
32	101	3.1	748	2	A45243	envelope protein HrpH	8.50e+00
33	102	3.1	756	2	C64236	protein V (fcrv) homo	6.65e+00
34	102	3.1	899	2	A44879	retinolastoma protel	6.65e+00
35	101	3.1	1123	1	WMBEH7	UL37 protein - human	8.50e+00
36	101	3.1	5762	2	A41819	proline-rich peptides	8.50e+00
37	99	3.0	218	2	F64461	hypothetical protein	1.38e+01
38	99	3.0	235	1	DNPDPW	repeat element protei	1.38e+01
39	99	3.0	259	2	B69113	cell division inhibit	1.38e+01
40	99	3.0	405	2	D70505	probable HflX - Mycob	1.38e+01
41	99	3.0	505	2	S09638	flagellin - Salmonell	1.38e+01
42	100	3.0	523	2	S46720	NAM6 protein - yeast	1.08e+01
43	99	3.0	527	2	A36175	amine oxidase (flavin	1.38e+01
44	100	3.0	849	2	JU0191	lysyl aminopeptidase	1.08e+01
45	100	3.0	1035	2	A64686	cation efflux system	1.08e+01

ALIGNMENTS

RESULT 1
 ENTRY A28568 #type complete
 TITLE pneumolysin - Streptococcus pneumoniae
 ORGANISM #formal name Streptococcus pneumoniae
 DATE 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change

ACCESSIONS A28568; S12829

REFERENCE A28568

#authors Walker, J.A.; Allen, R.L.; Falmagne, P.; Johnson, M.K.;

#journal Bouldin, G.J.

#title Infect. Immun. (1987) 55:1184-1189 -

Molecular cloning, characterization, and complete nucleotide sequence of the gene for pneumolysin, the

sulphydryl-activated toxin of Streptococcus pneumoniae.

#cross-references MIMD:87193109

#accession A28568

#molecule_type DNA

#residues 1-471 #label WAL

REFERENCE S12829

#authors Mitchell, T.J.; Mendez, F.; Paton, J.C.; Andrew, P.W.;

#journal Bouldin, G.J.

#title Nucleic Acids Res. (1990) 18:44010

Comparison of pneumolysin genes and proteins from

Streptococcus pneumoniae types 1 and 2.

#cross-references MIMD:90326546

#accession S12829

#molecule_type DNA

#residues 1-471 #label MIT

CLASSIFICATION #superfamily dipeptide transport protein

FEATURE #cross-references EMBL:X52474; NID:947403; PID:947404

2-471 #product pneumolysin #status predicted #label MAT

SUMMARY #length 471 #molecular-weight 52899 #checksum 8274

Query Match 99.9%; Score 3297; DB 2; Length 471;

Best Local Similarity 99.8%; Pred. No. 0.00e+00;

Matches 470; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB	1	MANKAVNDFILAMNYDKKLITLHGESEINRFTKEGNQDPDEEVIERKRSISTNTSDI	60
QY	1	MANKAVNDFILAMNYDKKLITLHGESEINRFTKEGNQDPDEEVIERKRSISTNTSDI	60
DB	61	SVTATNDSRLPGALLVVDLELLENNPTLLAVDRAPTYSIDPGLASSPSFLQVEDPSN	120
QY	61	SVTATNDSRLPGALLVVDLELLENNPTLLAVDRAPTYSIDPGLASSPSFLQVEDPSN	120

Db	121	SSVRCAVNDLLAKMHODYGOVNNP	PARMOYERKITAHSMEOLKYVFGSDFEKTNGLDIDF	160
Qy	121	SSVRCAVNDLLAKMHODYGOVNNP	PARMOYERKITAHSMEOLKYVFGSDFEKTNGLDIDF	160
Db	181	NSVHSGEKOIQIVNKKOITYTVSDAVNPNPDVODITYVEDLKORGISAEPLVYITSSV	240	
Qy	181	NSVHSGEKOIQIVNKKOITYTVSDAVNPNPDVODITYVEDLKORGISAEPLVYITSSV	240	
Db	241	AYGRQVYIKLETTKSDEVEAAFEALIKGVAVQADOTEMKOILLNTEYKAVILGGDPSSGA	300	
Qy	241	AYGRQVYIKLETTKSDEVEAAFEALIKGVAVQADOTEMKOILLNTEYKAVILGGDPSSGA	300	
Db	301	RYVYNGKVMVWEDLLOEGSRFTADRHGPISTTTSFRLDNNVATFQNSTDYETKVTAYRN	360	
Qy	301	RYVYNGKVMVWEDLLOEGSRFTADRHGPISTTTSFRLDNNVATFQNSTDYETKVTAYRN	360	
Db	361	GDLLLDHSGAVVAOYIYIMWDELSTVDHOCKEVLFPKAMDNRNGODLTAFHTSIPLKGVARN	420	
Qy	361	GDLLLDHSGAVVAOYIYIMWDELSTVDHOCKEVLFPKAMDNRNGODLTAFHTSIPLKGVARN	420	
Db	421	LSVKIRECTGLAMEWMTRVYERKTLPLVYRKRTTISIMGTTIYPOVEDYEVEND	471	
Qy	421	LSVKIRECTGLAMEWMTRVYERKTLPLVYRKRTTISIMGTTIYPOVEDYEVEND	471	

RESULT	2
ENTRY	B43577
TITLE	peritrypsin O precursor - Clostridium perfringens
ORGANISM	#formal_name Clostridium perfringens
DATE	03-Feb-1993 #sequence_revision 03-Mar-1993 #text_change
ACCESSIONS	18-Sep-1998 B43577; A34951; A60922

ACCESSIONS B43577, A34951, A60922
 REFERENCE A43577
 #authors Shimizu, T.; Okabe, A.; Minami, J.; Hayashi, H.
 #journal Infect. Immun. (1991) 59:137-142
 #title An upstream regulatory sequence stimulates expression of the perfringolysin O gene of *Clostridium perfringens*.
 #cross-references M01D:9109951

```
##molecule_type DNA
##residues 1-500 ##label SH1
##cross-references GB:M81080; NID:g144885; PID:g144886
##note translation of the nucleotide sequence is not completed
A194951
#####
```

#authors Tweten, R.K.
#journal Infect. Immun. (1988) 56:3235-3240
#title Nucleotide sequence of the gene for perfringolysin O (theta-toxin) from *Clostridium perfringens*: significant homology with the genes for streptolysin O and pneumolysin
#cross-references PMID:8903263

```
#molecule_type DNA
#residues 1-125,'EA',129-500 #label TWE
#cross-references GB:M36704; NID:g144883;
REFERENCE A60922
```

#Journal	Infect. Immun. (1988) 56:3228-3334
#Title	Cloning and expression in <i>Escherichia coli</i> of the petrinoglysin O (theta-toxin) gene from <i>Clostridium</i> <i>perfringens</i> and characterization of the gene product

```

#molecule_type protein
#residues 29-45 ##label TW2
#experimental_source ATCC 13124
GENETICS

```

```
#gene      pfo; pfoA
FEATURE
```

```

1-28      #domain signal sequence #status predicted #label SIGV
29-500    #product perfringolysin O #status experimental #label

```

```
SUMMARY
#length 500 #molecular-weight 55799 #checksum 4821
```

Query Match	49.0%;	Score 1618;	DB 2;	Length 500;
Best Local Similarity	46.9%;	Pred. No. 1.09e-272;		
Matches	216;	Conservative	110;	Mismatches 132;
			Indels	3;
			Gaps	3;

Db 35 NQSIDSGISSLYNNEVLASNGDKIES-FVPKEGKAGNKFIIVERQKRLTSPVDIS 93

QY 3 NKAIVDFILAMNYDKKKLLTHQGESIENRFI-KEGNQLPDEFVIERKKRSLSTNTSDIS 61

```

D6  IIDSVDNRITPGALQLDADKAFVEHRPTILMWKRRPILNINIDLPGLGKENS-IKVDDPTYG 152
94  :  :  |  |||  :  |  ::|  |  :  |  :  :  |||  :  :  |  :  :  :

```

02 VAINDSKILI FGALLV VDEI LHENNF I LLAVDRAFM I SIDPGLASSDSE LQVEDPSNS 121

122 SWINGWANT I KATHOLIKOVI NIMAS DIOKLETUS MURDERER OF THE EMPEROR 21

213 AVANNEKKVMTIAYKOTEVVSADI BK NPSDI EDDSVMENDI KOKCVSNEA BPI MYCMTA 27

QY 182 SVHSGEKOIOIVNEKYOITTVTSVDVKNPGDYEDOPTVVEDIKORGTSAERPI.VYIISVA 24

Db 273 YGRITVVKLETTSSSKDVAFAKALIKNTDIKNSQYKDIYENSSTAVVLGGDAQEHNK 33

QY 242 YGRQVYIKLETTSKSDYEAAFEALIKGVKVAQPTEWKQILDNTEVKAVILIGDPSSGAR 30

Db 333 VTKDEDIRKVNATSTKNPAPISYTSVEFKDNSVAAVHNKTDYIETTETYSKG 39

QY 302 VVTGKVDVEDLLQEGSHPTADHPGLPISTYTSFLDNVVAIFQNSTDYVEIKVTAIRNG 36

```
Db      393 KINLDHSGAYVAQFEVAMDEVSYDKEGNEVLTHKTDWGNYYQDKTAHYSTVIPLEANRNI 45
```

362 DLUDDHSGAVVAQYITINNELSYDHQKREVLTPKAMDNRNGQDLTAHFTTSIPLKGNVRL 422

00 ANANECIGLHMEWNRKVLSLDVPLINNINVSIMGILIP 493
00 :| | | | | | | | | | | | :| | : :| | | | | |
00 :| | | | | | | | | | | | :| | : :| | | | | |

402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

RESULT	3	A43505	#tune complete
ENTRY			

TITLE	ORGANISM
listeriolysin O precursor - <i>Listeria monocytogenes</i>	<i>Listeria monocytogenes</i>
#formal name	<i>Listeria monocytogenes</i>

```
DATE      21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change
17-Mar-1999
```

ACCESSIONS
 A43505; S05306; A4/606; S12400; A610/9
 REFERENCE
 A43505

Geoffroy, C.; Gicquel-Sanzey, B.; Baquero, F.; Perez-D-

#journal Infect. Immun. (1988) 56:766-772

listeriolysin O determinant of *Listeria monocytogenes*.
#cross-references MIMD:88153053

```
#accession      A43505
##status        preliminary
```

```
##molecule_type DNA
##residues 1-529 ##label MEN
```

```
##cross references GB:M24199; MID:g149052; PID:g149053
##note      this sequence is derived from a strongly hemolytic
```

REFERENCE	DATE	BY
S05306	1990-10-10	W. J. Chabrebert, E. Doman

#journal Nucleic Acids Res. (1989) 17:6406
#title Nucleotide sequence of the *hst1* gene from a *Yeast*

monocytogenes serotype 1/2a strain.
#cross-references MUID:89366684

```
#accession S05306
#molecule_type DNA
```

```

#RESIDUES      1-329 #1ADE1 DOM
##CROSS-references EMBL:X15127; NID:g44106; PID:g44107

```

```

#note      this sequence is derived from a weakly hemolytic strain
#sequence 1/2

```

REFERENCE
A47606

Query Match	Best Local Similarity	Matches	Conservative	Score	DB 2	Length	529
61	IDKTYQGLDYNKNNLVYHGDVAVTVNPPRKGKDGNGEYIVVEKKKKSINQNNADIQYVNA	47.2%	202	1558	DB 2	529	
6	VNDFTILAMNYDKKKLLTHQGESIERNFKEGNQLPDEFVIERKKRSLSTWTSIDSVAT	43.3%	202	1558	DB 2	529	
121	ISSLTYPGALVANSDELVENQDVPVPRKDSITLSDIPGNTNODKNIVYNNARKSNVNN	43.3%	202	1558	DB 2	529	
66	NDSRLYPGALVAVDDTLLENNPTLLAVRAPPTYSIDIDPGASSDSFLQVEDPSSSVRG	43.3%	202	1558	DB 2	529	
181	AVNTLYVERAKNEYAQAAPVPSAKIDYDDEMASESOLIAKFGTAFKAVNNSLVNFGAIS	43.3%	202	1558	DB 2	529	
126	AVNDLIAKHQYGOVN-NVPRKMYEKTITAHSMQLVKYKGSDEKTSIDIDENSVAH	43.3%	202	1558	DB 2	529	
241	EGRKQOEIVSFKOIYVNVNNEPTPRSPREFKAVTKEQLQALGVNAENPPAYISSVAYGR	43.3%	202	1558	DB 2	529	
185	SEKQIQIYNFQIYTVTVSVAADKKNPDGVQDTVEDLKQRGISAEPRIVYISSVAYGR	43.3%	202	1558	DB 2	529	
301	QVYLKISTNSHSTKKRAADAAVSGKSVSGVELTNIINKNSFKAVIYIGSKADEVQITD	43.3%	202	1558	DB 2	529	
245	QVYLKLEFTSKSDEVEAEALIKGVKAPQTEMKQIDNTEVKAIVILIGDPPSSGARVAT	43.3%	202	1558	DB 2	529	
361	GNLGDRLILKGAFTNRETPVPIATYTNPLKQDELAVIKNSSEYFTTSKAYTDGKIN	43.3%	202	1558	DB 2	529	
305	GKDYMEDILQGSRTFADHPCLPISYTSFSLRDVNVATFOUSTDYVEKVTAYAKMDL	43.3%	202	1558	DB 2	529	
421	IDHSGVYQAFINISMDENVYDDEGEIYQHKWMSNNKSLAHFTSSITLPGNANINIV	43.3%	202	1558	DB 2	529	
365	LDSHSGVAYQIYITINELSYDHQKEVLIPKAMDNGODLTAHFTTSLPLKGNVANSVYK	43.3%	202	1558	DB 2	529	
481	AKECTGLAEMWRTVIDRNDPLVKNRNTSINGTTLIPKYSKAKVDN	43.3%	202	1558	DB 2	529	
425	IRECTGLAEMWRTVEKTDLPVLRKRTISINGTTLIPVQVEDKVEN	43.3%	202	1558	DB 2	529	

[illegible]

ACCESSIONS	S22341; S36683
REFERENCE	S22340
#authors	Haas, A.; Dumbeky, M.; Kreft, J.
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Istiolysin genes: complete sequence of <i>ilo</i> from <i>Listeria</i>
##cross-references	Ivanovii and of <i>ilo</i> from <i>Listeria seeligeri</i> .
##accession	MUID:92182018
##molecule_type	DNA
##residues	52341
##cross-references	EMBL:X60461
##note	the authors translated the codon ACA for residue 331 as Val
REFERENCE	S36683
#authors	Kreft, J.
#submission	submitted to the EMBL Data Library, July 1991
#accession	S36683
#molecule_type	DNA
#residues	1-319,'T',321-528 ##label KRE
##cross-references	EMBL:X60461
GENETICS	
#gene	<i>ilo</i>
CLASSIFICATION	#superfamily dipeptide transport protein
FEATURE	
1-24	
25-528	
SUMMARY	
Query Match	46.7%; Score 1541; DB 2; Length 528;
Best Local Similarity	43.3%; Pred.No.5,70e-258;
Matches 207; Conservative 121; Mismatches 142; Indels 1; Gaps 1;	
Db	60 IDQYIGLDYDKNNILVVDGSAVKWVPKAKYKGNQYIVVEKKRSINONNADIQVNS 119
QY	6 VDNFLLANNYYDKRKLTHOGESIERFIKEGNQLPDERVVIIEKRKRSLSTNTSDISTAT 65
Db	120 LASLTYPALVKANSELVENOPDLVPYKRDSVTLSIDLPKVNHDEIIVQNATKSIND 179
QY	66 NDSRLYPPALLVDETLENNPTLLADVRAPIYSIDLPGIASDSPLOVEDPSNSSVRG 125
Db	180 GVNTLVDMNNKYSEHEYNISAKIIDYOEMAYSESOLVAKGAFKAVNNSLAWNFGAIS 239
QY	126 AVNDLLAKWHQDYGQVN NVPARROYEITHTSHSQGLKAYFGSPFEETGMSLIDERSVH 184
Db	240 EGKVOEEVINRKQIIYYVNVNEPSSPKRFEGSVTKENLCALGVANAENPRAYISVAAGR 299
QY	185 SGEOIOIVNRKQIIYYTVSYDAVNPBGVDYEDTLKORGISABRPITYISSVAYGR 244
Db	300 DIFVYLSTSSHSTRKAFAFDAPFKGSYKGDTELENLIIONASFRAYVYGGASDAVEIID 359
QY	245 QVYLKLETTSDEVEAFAELIIGVVYVAPDTWKQILDNTEKAVAILGGDPSSGARVYT 304
Db	360 GDLSLRDIILKOGANFNDRKNNGVBIATYTNFLKNOQLAAVKNNSEYITEISKAYSQGIN 419
QY	305 GKVMVEDLLIOEGSRFPADHPGLPISTTSPFLRDYNVAFTQNSDYETKVTAYRNQDLL 364
Db	420 LDHGGAVARARNVWDVSVDANSENVEYERKKSSENDKCLAHFTTISIYLGARAINIH 479
QY	365 LDHSGAIYAQQIYTWNELSIDHOKEVLPTRAMDRNQDDTAHFHTTISIPLKGVNRLSVK 424
Db	480 AKECTGLAMEWRVTVDNRNLPLVKNRNVCIWGTTLYPAYSDPVDN 525.
QY	425 IRECTGLAMEWRVYEKTDLPVYRKRTISIWGITTYLPVEDKRYEN 470
RESULT	6
ENTRY	S22340 #type complete
TITLE	seeligeriolysin - <i>Listeria seeligeri</i>
ORGANISM	#formal name <i>Listeria seeligeri</i>
DATE	22-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 21-Aug-1998
ACCESSIONS	S22340
REFERENCE	S22340

#authors	Haa, A.; Dumbsky, M.; Kreft, J
#journal	Biochim. Biophys. Acta (1992) 1130:81-84
#title	Listeriolysin genes: complete sequence of llo from <i>Listeria</i> Ivanovii and of lso from <i>Listeria seeligeri</i> .
#cross-references	M01D:92182018
#accession	S22340
#status	Preliminary
#molecule_type	DNA
##residues	1-530 ##label HAA
##cross-references	EMBL:X60462; NID:g44144; PID:g44145
##note	The authors translated the codon GCC for residue 287 as Pro
CLASSIFICATION	#superfamily dipeptide transport protein
SUMMARY	#length 530 #molecular-weight 59181 #checksum 4523
Query Match	45.5%; Score 1502; DB 2; Length 530;
Best Local Similarity	43.18; Pied. No. 1,59e-250;
Matches	201; Conservative 115; Mismatches 149; Indels 1; Gaps 1
Db	62 INKYLINGLNDKNSILYOGEAATVNPPEKKGKDCSEYIVEEKKKKGINONNADISVINA 121
Qy	6 VNDFILANWYDKKLLHGHSEIENPFIEGNDQDDEVIYERKKRSLSTSDISTAT 65
Db	122 ISSLTYPGALYKANRELVENQPVNLYPKKSDLSLTVDLPMTKKDKNKIYKKNPTKSNVN 181
Qy	66 NDSRLYPALLVDELTLENNPITLAVDRAPMYVSIIDPLGSSDSDFLOVEDDPNSSVRG 125
Db	182 AVNTLVEWMDKRSKAPNPJNAKIDVSDMAVSESLAKGTAFKAVNNSLVNPEAIS 241
Qy	126 AVNDLKAKHMDQGYNA-NVPAWQYKERTKTAISMQLKVGKSGDEFTGNSLDIDFNSVA 184
Db	242 DGYQVEYISFKQIYVINYINNEPTSPSKFEFGSVTRKQDLAAYNAENPPAYISSVAYGR 301
Qy	185 SGEQOIDIVNFKQIYYTSVSDAVKKNQGDVFOPTVIEDLKORGISAREPLVISSVAYGR 244
Db	302 QVYKLSSSSKNKKVAFAPFAAMSGKSVGVDLELNIIKNSFKAVIYGSAAEEVFIID 361
Qy	245 QVYIAKLETTKSDVEAFEPALIKGKYPAPQYEMKQILDNTEVAIVILGGDPSGARGVAT 304
Db	362 GNIGELBDILKSGSTYDRNPQVPISTYTNFKDNLDLAVVANNSESYIETTSKSYTGAKIN 421
Qy	305 GKVMVEDLDIQEGSRFLADHPGLPISTYTSFLRDVAVATPQNSIDYETKVTAVYRNGDLL 364
Db	422 IDHSGGVAAQFINSNDVSYDENGEMIKYHKWKGENTYSKLAHTSSIYLPGARINITY 481
Qy	365 LDHSGAAVAAQYITWELSYDQGEKVELYPRKAMPKRGODLTAHFTTIPLPKGVNRMLSVK 424
Db	482 ARECTGLFEMWARTVIDDRNLPLVKNRNVSLMGTLVPRHSNNVDN 527
Qy	425 IRECTGLAMEWMTYVEKTDLPVYKRTIISMGTLIPQVEDKVEN 470
RESULT	7
ENTRY	A37858 #type complete
TITLE	alveolysin - <i>Bacillus alvei</i>
ORGANISM	#formal name <i>Bacillus alvei</i>
DATE	28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 09-Sep-1997
ACCESIONS	A37858
REFERENCE	A37858
#authors	Geoffroy, C.; Mengaud, J.; Alouf, J.E.; Cossart, P.
#journal	<i>J. Bacteriol.</i> (1990) 172:7301-7305
#title	Alveolysin, the thiol-activated toxin of <i>Bacillus alvei</i> , is homologous to listeriolysin O, pertissinolysin O, pneumolysin, and streptolysin O and contains a single cysteine.
#cross-references	M01D:91072294
#accession	A37858
##status	Preliminary
##molecule_type	DNA
##residues	1-501 ##label GEO
##cross-references	GB:W62079; NID:9142473; PID:g142473
SUMMARY	#length 501 #molecular-weight 55268 #checksum 8775

Db	402	TLDHGAVAADPDSMDQFTEEDONGKELTEHKWTMGSGKDKATAYVYPLPAPSKNIKI	461
Qy	364	LIDHSGAIVAAQYITIMWNLSTIDHSGKEVLTEPKAMDRNGQDLTAHFTTSTPLKGNVNLISV	423
Db	462	VARCTGLAMEWMRTI	477
Qy	424	KIRECTGLAMEWMRTV	439
RESULT	10		
ENTRY		S47298	#type complete
TITLE		su1ysin - Streptococcus suis	
ORGANISM		#formal_name Streptococcus suis	
DATE		06-Jan-1995	#sequence_revision 06-Jan-1995
ACCESSIONS		08-Sep-1997	#text_change
REFERENCE		S47298	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		S47298	
##status		preliminary	
##molecule_type		DNA	
##residues		1-112	#label SEG
##cross-references		EMBL:Z36908; NID:9535307; PID:9535308	
SUMMARY		#length 112	#molecular_weight 12464
			#checksum 9264
Query Match		14.1%; Score 464; DB 2; Length 112;	
Best Local Similarity		50.9%; Pred. No. 4.77e-57;	
Matches		57; Conservative 22; Mismatches 33; Indels 0; Gaps 0;	
Db	1	PGVPVSTSTFVXXXXXPRQXLSNSEYETFTTIVNNSSLTLXHSAGAVAKNTWGEVSY	60
Qy	325	PGLPSTYTSFLRDNVYATFONSTDYETKRAYRNGDLDLDHSGAIVAAQYITIMWNLST	384
Db	61	NEAGEEVEWEXKAMPDKNGVNLTSNHSSETIXIPGNAXNLHVNIQCTGLAMEWM	112
Qy	385	DHSGKEVLTEPKAMDRNGQDLTAHFTTSTPLKGNVNLISV	436
RESULT	11		
ENTRY		S47297	#type complete
TITLE		su1ysin - Streptococcus suis	
ORGANISM		#formal_name Streptococcus suis	
DATE		06-Jan-1995	#sequence_revision 06-Jan-1995
ACCESSIONS		09-Sep-1997	#text_change
REFERENCE		S47297	
#authors		Segers, R.	
#submission		submitted to the EMBL Data Library, September 1994	
#accession		S47297	
##status		preliminary	
##molecule_type		DNA	
##residues		1-96	#label SEG
##cross-references		EMBL:Z36907; NID:9535305; PID:9535306	
SUMMARY		#length 96	#molecular_weight 10471
			#checksum 6662
Query Match		6.8%; Score 225; DB 2; Length 96;	
Best Local Similarity		34.8%; Pred. No. 8.99e-17;	
Matches		21; Conservative 25; Mismatches 33; Indels 0; Gaps 0;	
Db	1	KODNVPQSLTYGQRLITNGEYIINDNPATTTGALENGREYVLRREKXITTDNSADITVI	60
Qy	4	KAVNDFILIAMYDKRKLTLTHGSEIENFNKIEGQQLPDEFVIERKKRSLSTNSDISVT	63
Db	61	XAKANISPGALLRAXOVLDDNNPTLISI	89
Qy	64	ATNDRSLYPGALLVYDETLLENPTLLAV	92
RESULT	12		
ENTRY		S03974	#type complete
TITLE		amine oxidase (flavin-containing) (EC 1.4.3.4) A - bovine	
ALTERNATE_NAMES		monoamine oxidase type A	

ORGANISM	#formal_name	Bos primigenius taurus	#common_name	cattle
DATE	28-Feb-1990	#sequence_revision	28-Feb-1990	#text_change
ACCESSIONS	503974			
REFERENCE	503974			
#authors	Powell, J.F.; Hsu, Y.P.P.; Weyler, W.; Chen, S.; Salach, J.;			
#journal	Biochem. J. (1989) 259:407-413			
#title	The primary structure of bovine monoamine oxidase type A. Comparison with peptide sequences of bovine monoamine oxidase type B and other flavoenzymes.			
#cross-references	MUID:89246344			
#accession	S03974			
#status	not compared with conceptual translation			
#molecule_type	mRNA			
#residues	1-527	#label	POW	
#cross-references	GB:X15609; NID:g5523; PID:g5524			
KEYWORDS	FAD; flavoprotein; mitochondrion; oxidoreductase			
FEATURE				
15-43				
SUMMARY	#length 527	#molecular_weight	59800	#checksum 6598
Query Match	3.8%;	Score	124;	DB 2; Length 527;
Best Local Similarity	20.0%;	Pred. No.	1.95e-02;	
Matches	30;	Conservative	42;	Mismatches 72; Indels 6; Gaps 5;
Db	1	MESLQKTSDAQMDVVYVIGGIGISGLSAAR-LIAEHVNVNLVEARRRGGRITYANEH	59	
Oy	250	LETTSKSDEVAEALFKGVKAPDTEMKOILDNTVTKAVILGGDSSGARVVTGKVDM	309	
Db	60	V-DYVDVGAGVGTQNRILRLSKOLELTYKVNVERLHYVKGKTPPRGAFPPWNP	118	
Oy	310	VEDLIQESRTADHPGLPISTYSF-LRNVYATFPNSDIYETKTAARNDDLHDH	368	
Db	119	IATV-DYNNLMRTM-DNMGKEIPADAPWE	145	
Oy	369	GAYVAQYITWNELSYDHQKEVLETPKAWD	398	
RESULT	13			
ENTRY	T03415	#type	complete	
TITLE	S-layer protein - Campylobacter rectus			
ORGANISM	#formal_name	Campylobacter rectus		
DATE	16-Mar-1999	#sequence_revision	16-Mar-1999	#text_change
ACCESSIONS	T03415			
REFERENCE	214942			
#authors	Wang, B.; Kraib, E.; Kolodrubetz, D.			
#journal	Infect. Immun. (1998) 66:1521-1526			
#title	A new member of the S-layer protein family: characterization of the crs gene from Campylobacter rectus.			
#accession	T03415			
#status	preliminary			
#residues	1-1361	#label	MAN	
#cross-references	EMBL:AF010143; NID:g2459960; PID:g2459961.			
SUMMARY	#length 1361	#molecular_weight	144385	#checksum 6731
Query Match	3.6%;	Score	118;	DB 3; Length 1361;
Best Local Similarity	22.0%;	Pred. No.	1.03e-01;	
Matches	41;	Conservative	49;	Mismatches 86; Indels 10; Gaps 10;
Db	178	VDWHSNPEHPKAELETTNDMAANFANPMK-HNNGGCDRLMTLOSSKLTGDSRHD	236	
Oy	115	VEDPSSNSVRCAVNDLAKMHODYGOVNNYPARQYEKIRAHSMK-QLKVKFSSDEKGTG	173	
Db	237	NTLVNEFGQANADGDEPTSRPTLTNTQININIEVTGVTNLDL-RDSNDYKINIRITK	295	
Oy	174	NSLDIDENSVAHSGE-KQI-QLVNFQKI-YTTVSVDANKNGDVPQDITVYEDLKQGISA	230	
Db	296	EAGKKFVNESIG-OKLVGMRLANYAKKD-IDVAFEH-KKCVLSGFEDKSNVPLENVEAKS	352	

```

QY 231 EEPYLV-ISSVAYGROYVTKLETTSSKDEVEAFALIKGVAVAQTEWKQILDVTEYKA 289
DB 353 LSTSP 358
QY 290 VILGGD 295

RESULT 14
ENTRY 560947 #type complete
TITLE hypothetical protein YOR220w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein O5050; hypothetical protein YOR50-10
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 05-Jun-1998
ACCESSIONS 560947; S67113; S71722
REFERENCE 560938
#authors Gallison, F.; Dujon, B.
#submission submitted to the EMBL Data Library, October 1995
#description Sequence and analysis of a 33 kb fragment from the right arm of Chromosome XV of the yeast Saccharomyces cerevisiae.
#accession S60947
##molecule_type DNA
##residues 1-265 #label GAL
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
REFERENCE S67104
#authors Boyer, J.; Fairhead, C.; Gallion, L.; Gallison, F.; Michaux, G.; Pierrry, A.; Dujon, B.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67113
##molecule_type DNA
##residues 1-265 #label BOY
#cross-references EMBL:Z75128; NID:g1420509; PID:e252085; PID:g1420510; MIRS:YOR220w
##experimental_source strain S288C
REFERENCE S71713
#authors Gallison, F.; Dujon, B.
#journal Yeast (1996) 12:877-885
#title Sequence and analysis of a 33 kb fragment from the right arm of Chromosome XV of the yeast Saccharomyces cerevisiae.
#cross-references MUID:96437977
#accession S71722
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-265 #label GAW
#cross-references EMBL:X92441; NID:g1050762; PID:g1050772
#note the nucleotide sequence was submitted to the EMBL Data Library, October 1995
GENETICS
#map_position 15R
SUMMARY #length 265 #molecular_weight 29255 #checksum 6028
Query Match 3.5%; Score 117; DB 2; Length 265;
Best Local Similarity 27.6%; Pred. No. 1.35e-01;
Matches 29; Conservative 28; Mismatches 41; Indels 7; Gaps 7;
DB 85 PMKLFVTESLNNO-HPR-SRSTDVAISLONNALLDHRNRKLSINDPVGTVGDS 142
QY 12 AMN-YDKKKLLTHGSEENRFKEGNOLPDE-FVYLE-RRKRSLSTNTSDISTATNDS 68
DB 143 SLNKGSSLSFDKSLSPMLKLSIDSPKPSYOEPLPKLSRSS 187
QY 69 RLY-PGALLVVDFTLLNNPTL-LAVDRAPMTYSIDLPGIASSDS 111

RESULT 15
ENTRY S58322 #type complete
ALTERNATE_NAMES nucleolar protein NOP5 - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 24-Sep-1998

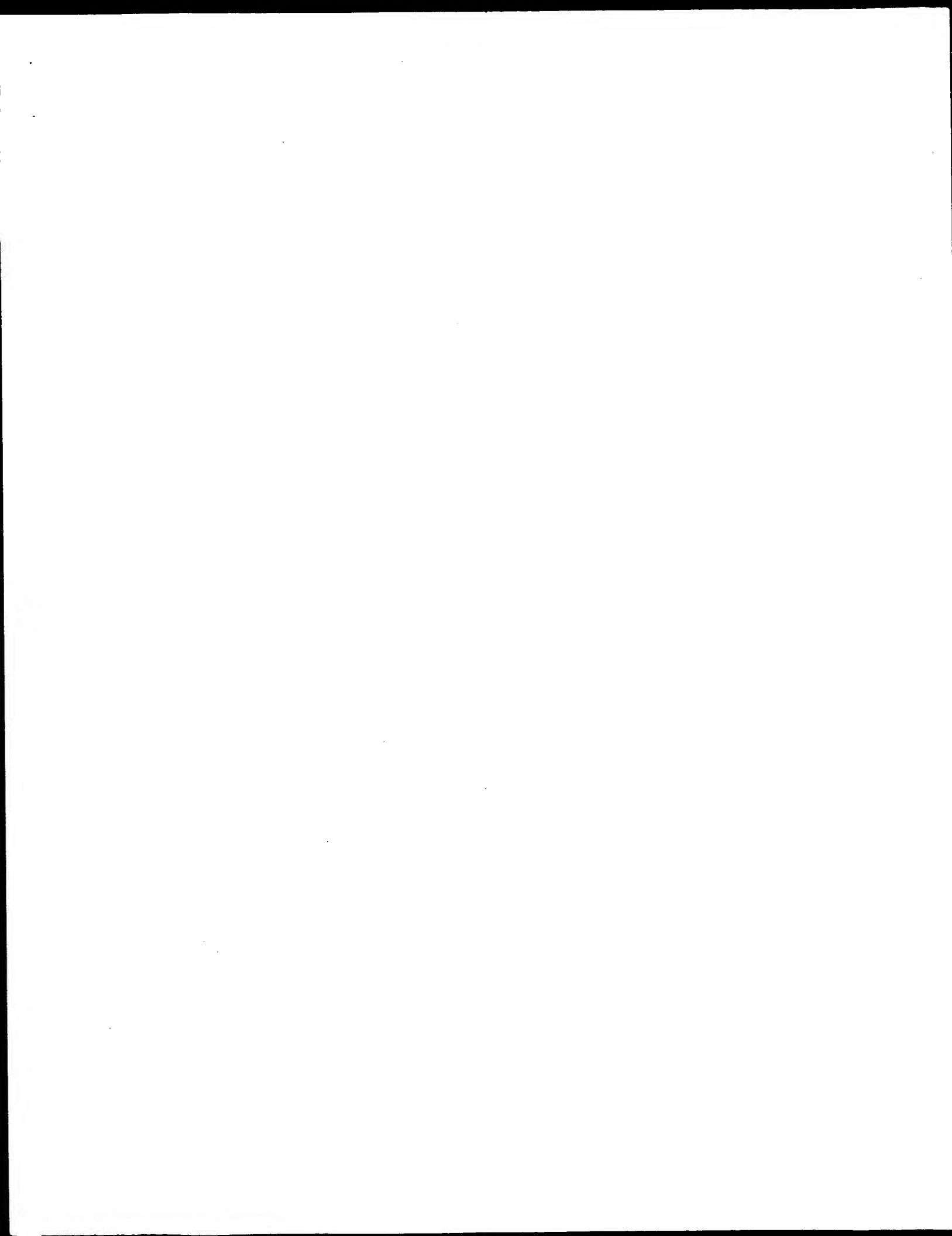
```

```

ACCESSIONS S58322; S67216; S71990
REFERENCE S58318
#authors Pearson, B.M.; Hernando, Y.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
#submission submitted to the EMBL Data Library, August 1995
#accession S58322
##molecule_type DNA
##residues 1-511 #label PEA
#cross-references EMBL:X90565; NID:g940836; PID:g940841
REFERENCE S67213
#authors Pearson, B.M.; Hernando, Y.; Kalogeropoulos, A.; Schweizer, M.
#submission submitted to the Protein Sequence Database, July 1996
#accession S67216
##molecule_type DNA
##residues 1-511 #label PEW
#cross-references EMBL:Z75217; NID:g1420680; PID:e252143; PID:g1420682; MIRS:YOR310C
##experimental_source strain S288C
REFERENCE S71986
#authors Pearson, B.M.; Hernando, Y.; Payne, J.; Wolf, S.S.; Kalogeropoulos, A.; Schweizer, M.
#journal Yeast (1996) 12:1021-1031
#title Sequencing of a 35.71 kb DNA segment on the right arm of Chromosomes I and XIII.
#accession S71990
#status nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues 1-511 #label PER
#cross-references EMBL:X90565; NID:g940836; PID:g940841
#note the nucleotide sequence was submitted to the EMBL Data Library, August 1995
GENETICS
#gene SGD:NOP5
#cross-references SGD:S0005837; MIRS:YOR310C
#map_position 15R
FUNCTION #description involved in the synthesis of the 40S ribosomal subunit
KEYWORDS nucleus
SUMMARY #length 511 #molecular_weight 56956 #checksum 4899
Query Match 3.4%; Score 112; DB 2; Length 511;
Best Local Similarity 22.8%; Pred. No. 5.16e-01;
Matches 31; Conservative 39; Mismatches 58; Indels 8; Gaps 8;
DB 188 YGMH-FPELAKIV-TDSVAYARIITLTMGRSKASPTDISEILPEIEER-VKTAAEVSMG 244
QY 242 YGRQVYTKLETTSSKDEVEAFALIKGVAVAPQTEWKQILDVTEKAVILGGDPSSG 299
DB 245 TEITQTDIDNINALAEQIVFAAYREOLS-NYLSARMAKAIAPNLTQVGEIVGARLII-H 302
QY 300 ARVYVGKVDVMDLEDIOGSRFTADHPGLPISTYSFLDNDVAVTFQNSTDVEIKVAYR 359
DB 303 SGLSLSLAKSPASTIQ 318
QY 360 NGDLL-LDHSAGVVAQ 374

Search completed: Mon Aug 30 16:08:05 1999
Job time : 59 secs.

```

 WIRE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 16:04:38 1999; Maspar time 14.67 Seconds

Tabular output not generated. 907.352 Million cell updates/sec

Title: >US-09-120-044-3

Description: (1-471) from US09120044.pep

Sequence: 1 MANKAVNDFILANMYDKKL.....TISWGCTTLYPYQEVKEND 471

Scoring table: PAM 150

Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot37

1:swissprot

Statistics: Mean 51.760; Variance 98.133; scale 0.527

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	3286	99.6	470	1	TACY-STPRN	PNEUMOLYSIN (THIOL-ACT	0.00e+00
2	1618	49.0	500	1	TACY-CLOPE	PERFRINGOLYSIN O PRECU	0.00e+00
3	1558	47.2	529	1	TACY-LISMO	LISTERIOLYSIN O PRECU	4.39e-303
4	1539	46.7	528	1	TACY-LISIV	IVANOLYSIN PRECURSOR (7.16e-299
5	1502	45.5	530	1	TACY-LISSE	SEELIGERIOLYSIN PRECU	1.13e-290
6	1492	45.2	501	1	TACY-PAEAL	ALVEOLYSIN PRECURSOR (1.85e-288
7	1437	43.6	574	1	TACY-STREO	STREPTOLYSIN O PRECURS	2.74e-276
8	1436	43.5	571	1	TACY-STREY	STREPTOLYSIN O PRECURS	4.56e-276
9	1430	43.0	574	1	TACY-STREY	STREPTOLYSIN O PRECURS	1.57e-272
10	1388	41.2	485	1	TACY-STREB	HEMOLYSIN PRECURSOR (F	7.75e-259
11	1348	39.8	527	1	TACY-BACCE	AMINE OXIDASE [FLAVIN-	1.09e-01
12	1112	34.4	511	1	NOPE-YEAST	NUCLEOLAR PROTEIN NOPE	1.97e-01
13	1110	33.3	872	1	DP3A-MYCPN	DNA POLYMERASE III, AL	1.97e-01
14	110	3.3	1228	1	SLAP-BACST	S-LAYER PROTEIN PRECU	1.97e-01
15	105	3.2	146	1	CASK-UNCUN	KAPPA CASEIN (FRAGMENT	6.42e-01
16	106	3.2	462	1	MPFB-YEAST	MITOCHONDRIAL PROCEST	6.33e-01
17	102	3.1	418	1	Y285-HELPY	HYPOTHETICAL PROTEIN H	2.58e+00
18	101	3.1	427	1	GLYA-SYNY3	SERINE HYDROXYMETHYLTR	2.58e+00
19	103	3.1	512	1	ANX7-XENLA	ANNEKIN VII (SYNEXIN)	1.48e+00
20	101	3.1	748	1	HRPH-PSESY	HYPERSENSITIVITY RESO	2.58e+00
21	102	3.1	756	1	V328-MYCGE	HYPOTHETICAL PROTEIN M	1.56e+00
22	101	3.1	1123	1	V120-HSV11	CAPSID ASSEMBLY PROTEI	2.58e+00
23	98	3.0	210	1	SOMA-ONCKI	SOMATOTROPIN PRECURSOR	5.84e+00

24	99	3.0	235	1	REEP-CSV	REPEAT ELEMENT PROTEIN	4.46e+00
25	99	3.0	350	1	ARGI-SOYBN	ARGINASE (EC 3.5.3.1).	4.46e+00
26	99	3.0	504	1	FLIC-SALMO	FLAGELLIN (PHASE-1-D F	4.46e+00
27	98	3.0	513	1	YJIV-ECOLI	HYPOHETICAL 58.2 KD P	5.84e+00
28	100	3.0	523	1	NAM8-YEAST	NAM8 PROTEIN.	3.40e+00
29	99	3.0	527	1	AOFA-HUMAN	AMINE OXIDASE [FLAVIN-	4.46e+00
30	99	3.0	723	1	HS9B-MOUSE	HEAT SHOCK PROTEIN HSP	4.46e+00
31	100	3.0	848	1	AMPN-LACTA	AMINOPEPTIDASE N (EC 3	4.46e+00
32	99	3.0	993	1	VIA-CMT11	1A PROTEIN (CONTAINS:	4.46e+00
33	98	3.0	1026	1	BGAL-STRTA	BETA-GALACTOSIDASE (EC	5.84e+00
34	95	2.9	177	1	GCSE-YEAST	GLYCINE CLEAVAGE SYSTE	1.29e+01
35	97	2.9	364	1	GCH2-PROLE	GTP CYCLOHYDROLASE II	7.63e+00
36	96	2.9	402	1	ARGJ-METJA	PUTATIVE GLUTAMATE-N-A	9.94e+00
37	97	2.9	419	1	V223-MYCPN	HYPOTHETICAL PROTEIN M	7.63e+00
38	96	2.9	527	1	IF2G-YEAST	EUKARYOTIC TRANSLATION	9.94e+00
39	96	2.9	662	1	LOXE-MOUSE	ARACHIDONATE 12-LIPOXY	9.94e+00
40	97	2.9	722	1	YG42-YEAST	HYPOTHETICAL 78.8 KD P	7.63e+00
41	96	2.9	767	1	HYPF-SYNY3	TRANSCRIPTIONAL REGUL	9.94e+00
42	97	2.9	1235	1	CY44-TRYBB	RECEPTOR-TYPE ADENYLAT	7.63e+00
43	97	2.9	1390	1	RPOB-MYCGE	DNA-DIRECTED RNA POLYM	7.63e+00
44	97	2.9	1675	1	CLH-RAT	CLAIRIN HEAVY CHAIN.	7.63e+00
45	96	2.9	3027	1	POLG-PYFV1	GENOME POLYPROTEIN [CO	9.94e+00

ALIGNMENTS

RESULT 1
 ID TACY-STPRN STANDARD: PRT: 470 AA.
 AC P11990;
 DT 01-OCT-1989 (REL. 12, CREATED)
 DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE PNEUMOLYSIN (THIOL-ACTIVATED CYTOLYSIN).
 GN PLY.
 OS STREPTOCOCCUS PNEUMONIAE.
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
 CC STREPTOCOCCUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCC 7466 SEROTYPE II;
 RX MEDLINE; 87193109.
 RA WALKER J.A., ALLEN R.L., FALMAGNE P., JOHNSON M.K., BOUNOIS G.J.;
 RT "molecular cloning, characterization, and complete nucleotide
 sequence of the gene for pneumolysin, the sulfhydryl-activated toxin
 of Streptococcus pneumoniae".
 RL INFECT. IMMUN. 55:1184-1189(1987).
 CC - FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 EUKARYOTIC CELL MEMBRANES.
 CC - SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X52474; G47404; -.
 CC EMBL: M17717; G153692; -.
 CC PIR: A28566; A28568.
 CC PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
 CC PFAM: PF01289; THIOL_cytolysin; 1.
 CC HSSP: P19995; LPFO.
 CC TOXIN; HEMOLYSIS; CYTOLYSIS.
 FT INIT MET 0
 FT ACT_SITE 427 427 BINDING TO CHOLESTEROL (BY SIMILARITY).
 SQ SEQUENCE 470 AA; 52768 MW; D3F3A252 CRC32;
 Query Match 99.6%; Score 3286; DB 1; Length 470;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;

	Matches	469;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps
Dd	1	ANKAVNDFLIAMVYDKKKLLTHQGSINERIRIKGNGLPDEFFVIERRKRLSTNTSDIS	60						
Oy	2	ANNAVDFILIAMWDRKKLTLHQGESIENRIKRKGNNLPDEFVIERRRKSLSLTNSTDIS	61						
Dd	61	VTAANDSRLEPGALLVDETTELENPTLLADRAPMTYSIDLEGLASDSFLOVEDPPNS`	120						
Oy	62	VTTATNRSRLYPGALLVDVEETELLENPITLLAADRAPMRTYSISDLEGASDSOSFLOVEDPPNS	121						
Dd	121	SVRGA VND LLA KHOD YGO VNV N PARM OYE K I T A H S M Q LKV R G SD E FKT G NS LD IDFN	180						
Oy	122	SVRGAVNDLLAKHODYGVGVNVVPARMOEKITAHSMEOQLVKFGSFEFTGTNSLDIDFN	181						
Dd	181	SYSHGEKOIOYNFKOIIYYTVSVDAVKRNGVFDDDTVTEDLKORGISAREPLVYTSSVA	240						
Oy	182	SYSHGEKOIOYNFKOIITYTSVDNAKNRPGDVFDOTVTEDLKORGISAERP LVYTSSYA	241						
Dd	241	YGROYVLKETLETSSKSDVENAEAFALIKGVAPOTEWKQILDNTWEKYAILGGDPSSGAR	300						
Oy	242	YGRVOYLKETLETSSKSEVEPAEFALKGVAPAOTEWKQILDNTWEKAYILGGDPSSGAR	301						
Dd	301	VVMGKVDMVEDLIQESSRETPAHPCLPISTTSLRNVAATONSNDYVERKVARNRG	360						
Oy	302	VVGKDVMDIEDLIQESRSFPADHPCLPISTTSLRNAVAATONSDYVERKVARNRG	361						
Dd	361	DLILDHSAGVAAYIYTMDELSDYDHQGKEVTFPRAMRNCGODTLAHTTISIPLKGNRTL	420						
Oy	362	DLILDHSAGVAAYIYTWNELSDYDHQREVLFPRAMBRNGODTLAHTTISIPLKGNVRTL	421						
Dd	421	SVKIRECTGLAMEWRMYEKEIDPLEYKRRITSIMTGTLIPOVEDKYEND	470						
Oy	422	SVKIRECTGLAMEWRMYEKTDPLEYKRRITSIWGTTLIPOVEDKYEND	471						
Pcscit_T_	2	TACY CLOPE	STANDARD:	PRT:	500 AA.				
AC	p19995,								
Dt	01-FEB-1991 (REL. 17, CREATED)								
Dt	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)								
Dt	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)								
De	PERFRINGOLYSIN O PRECURSOR (THETA-TOXIN) (THIOL-ACTIVATED CYTOLYSIN).								
Gn	PRO OR POR OR PCOA.								
Os	CLOSTRIDIUM PERFRINGENS.								
OC	PLASMID PRIB.								
CC	BACTERIA; FRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; CLOSTRIDIACEAE;								
NC	CLOSTRIDIUM.	[1]							
RN	SEQUENCE FROM N.A.								
RA	MEDLINE: 9109951.								
RA	SHIMITSU T., OKABE A., MINAMI J., HAYASHI H.: "An upstream regulatory sequence stimulates expression of the perfringolysin O gene of Clostridium perfringers."; INFECT. IMMUN. 59:137-142(1991). [2] [1] SEQUENCE FROM N.A., AND SEQUENCE OF 29-44. RP STRAIN=ATCC 13124; RC MEDLINE: 89032623. RX TWETTEN R.K.; RA "Nucleotide sequence of the gene for perfringolysin O (theta-toxin) RT from Clostridium perfringers: significant homology with the genes for RT streptolysin O and pneumolysin."; RM INFECT. IMMUN. 56:3235-3240(1988). [3] RN SEQUENCE OF 29-45 AND 305-312. RC STRAIN=PBGK; RX MEDLINE: 87076517. RA OHNO-IWASHITA Y., IWAMOTO M., MITSUI K., KAWASAKI H., ANDO S.; RT "Cold-labile hemolysin produced by limited proteolysis of theta-toxin RL from Clostridium perfringers"; RL BIOCHEMISTRY 25:6048-6053(1986). [4] RN SEQUENCE OF 492-500 FROM N.A.								

```

RX STRAIN-NOTC 8237;
RX MEDLINE: 96123363.
RA SHIMIZU T., KOHAYASHI T., BA-THEN W., OHTANI K., HAYASHI H.;
RT "Sequence analysis of flanking regions of the pfoA gene of
RT Clostridium perfringens: beta-galactosidase gene (pbp) is located in
RL the 3'-flanking region.";
RN MICROBIOL. IMMUNOL. 39:677-686(1995).
[5]
RP IMPORTANCE OF THIOL-GROUP IN CHOLESTEROL BINDING.
RX MEDLINE: 88004463.
RA IWAMOTO M., OHNO-IWASHITA Y., ANDO S.;
RT "Role of the essential thiol group in the thiol-activated cytolysin
RT from Clostridium perfringens.";
RN EUR. J. BIOCHEM. 167:425-430(1987).
[6]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTOMS).
RX MEDLINE: 97113505.
RA FELL S.C., ROSSJOHN J., RODE K., TWETEN R.K., PARKER M.W.;
RT "Crystallization and preliminary x-ray analysis of a thiol-activated
RT cytolysin.";
RN FEBS LETT. 397:290-292(1996).
[7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTOMS).
RX MEDLINE: 97325744.
RA ROSSJOHN J., FELL S.C., MCKINSTRY W.J., TWETEN R.K., PARKER M.W.;
RT "Structure of a cholesterol-binding, thiol-activated cytolysin and a
RT model of its membrane form.";
RN CELL 89:685-692(1997).
CC -! FUNCTION: SUPERNATANT-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL.
CC CONTRAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -! SUBUNIT: FORMS OLIGOMERS IN THE HOST MEMBRANE.
CC -! SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M81080; GI44886; -.
DR EMBL: M86704; GI44884; -.
DR EMBL: D49537; G1502275; -.
DR PIR: B43577; B43577.
DR PDB: 1PFO; 05-AUG-98.
DR PROSITE: PS00481; THIOL_CYTOLYSINS; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMD; 3D-STRUCTURE.
FT SIGNAL 1 28
FT CHAIN 29 500 PERFRINGOLYSIN O.
FT ACT_SITE 459 459 BINDING TO CHOLESTEROL.
FT CONFLICT 30 30 D-> K (IN REF. 3).
FT CONFLICT 34 34 K-> I (IN REF. 3).
FT CONFLICT 126 128 RKP-> EA (IN REF. 2).
SQ SEQUENCE 500 AA; 55799 MW; 17FC32BC CRC32;

```

```

QY      122 SVFGAANDLLAKHODVGGCVANNPARNQYEKIIAHSEOLKAYFSGPEFKTGSLLIDEN 181
Db      213 AYAANKKKWILLAYKQIFITYVSADLPKRNSDPLFDDSVTFENDLKOKGVSNAPLAWSNVA 272
QY      182 ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::| 241
QY      273 YKRTIYVKLETTSSSDVOAAAFKALKLNMDIKNSOQYKOIYESSFTAVVLGGDAOEHNK 332
QY      242 YGRQYVLEKIETTSKSDEVEAEFAALRKGVKVAPOTEWKRLDNTWEKAVYLGGDPSSGAR 301
Db      333 VYTKDELRKYIKDNATFSTKNPAYPISTYSVELKDNSVAAAHNKNDIETETSTESKG 392
QY      302 VYTGKVDMEYDLIQEGRSPTADHPGLISTTSFEFLDNVATQNSTDYIETVATYRNG 361
Db      393 KINLDHSGAYVAQEFVAMDEVSTDKGNEVLTHTKWGDNYQDTAAHYSTVPLEANARNI 452
QY      362 DLLLHSHGAYVAQYITLWELSYDHOGKEVLPRKAMDNRNQDLTAFTTISPLKGNVRNL 421
Db      453 RIKARECTGLAMEWRWDVISEVDYVPLTNININWSIMCTIYLP 493
QY      422 SYKIRECTGLAMEWRWTYTEKTDLPLVRRTTISIWGTYLP 462

RESULT
3
ID TACY LISTMO STANDARD: PRT: 529 AA.
AC PL3128: OA8747; O57096; O57206;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LISTERIOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN HLY OR LISA OR HLXA
OS LISTERIA MONOCYTOGENES.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
OC LISTERIA.
RN [1]
RN RP SEQUENCE FROM N.A.
RX STRAIN=EGD / SEROTYPE 1/2A;
RC MEDLINE: 89366684.
RA DOMANN E., CHARRABORTY T.;
RT "Nucleotide sequence of the listeriolysin gene from a Listeria
RT monocytogenes serotype 1/2a strain."
RL NUCLEIC ACIDS RES. 17:6406-6406(1989).
[2]
RV SEQUENCE FROM N.A.
RX MEDLINE: 88153053.
RA MENGARD J., VICENTE M.-F., CHENEVERT J., PEREIRA J.M., GEOFFROY C.,
RA GIQUEUET-SANZEB B., BAQUERO F., PEREZ-DIAZ J.-C., COSSART P.,
RT "Expression in Escherichia coli and sequence analysis of the
RT listeriolysin O determinant of Listeria monocytogenes.";
RL INECT. IMMUN. 56:766-772(1988).
[3]
RN RP SEQUENCE FROM N.A.
RX STRAIN=12067;
RC MEDLINE: 92040062.
RA RASMUSSEN O.F., BECK T., OLSEN J.E., DONS L., ROSSEN L.;
RT "Listeria monocytogenes isolates can be classified into two major
RT types according to the sequence of the listeriolysin gene.";
RL INECT. IMMUN. 59:3945-3951(1991).
[4]
RN RP SEQUENCE FROM N.A.
RX STRAIN=F265; F5782 (SEROTYPE 4B), AND F4233, F6789 (SEROTYPE 1/2B);
RC VINES A., SAMMINITHAN B.;
RA SUBMITTED (JUL-1995) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -I- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIMATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -I- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

[illegible]

```
OS      ILIO.
ON       LISTERIA IVANOVII.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACTILACEAE;
OC     LISTERIA.
RN   [1]
RP    SEQUENCE FROM N.A.
RC STRAIN-ATCC 19119;
RX MEDLINE; 92182018.
RA HAAS A.; DUMBESKY M., KREFT J.:
RT "listeriolsin genes: complete sequence of ilo from Listeria ivanovii
RL and of ilo from Listeria seeligeri."
BIOLCHIM. BIOPHYS. ACTA 1130:81-84(1992).
CC -I- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC ENKAROTIC CELL MEMBRANES.
CC -I- SIMILIARTY: BELONGS TO THE THIOL-ACTIVATED CYTOLOXSIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to licens@sib-slb.ch).
-----
DR EMBL; X60461; E39053; ALT INIT.
DR PROSITE; PS00481; THIOL_CYTOLOXSINS; 1.
DR PFAM; PF01289; Thiol_cytolysin; 1.
DR HSSP; P19995; lpro.
KW TOXIN; HEMOLYSIS; CYTOLOXSIS; SIGNAL.
FT SIGNAL                     1         23
FT CHAIN                      24        528
FT ACT_SITE                   483        483          BINDING TO CHOLESTEROL (BY SIMILIARTY).
SQ SEQUENCE                    528 AA; 58542 MW; CFAE84AD CRC32;

Query Match           46.7%; Score 1539; DB 1: Length 528;
Best Local Similarity 43.1%; Pred. No. 7,1be-299;
Matches 201; Conservative 122; Mismatches 142; Indels 1; Gaps 1;
```

	RESULT	5	STANDARD:	PRT:	530 AA.
AC	TACY_LISSE				
DR	P31830;				
DT	01-JUL-1993 (REL. 26, CREATED)				
DT	01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)				
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)				
DE	SERIALIZED TO PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).				
GN	LISO.				
OS	LISTERIA SEELIGERI.				
OC	BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;				
OC	LISTERIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SLOC;				
RX	MEDLINE; 92182018,				
RA	HAAS A., DOMESKY M., KREFT J.:				
RT	"Listeria lysin genes: complete sequence of llo from Listeria ivanovi				
RL	and of lso from Listeria seeligeri."				
RL	BIOCHIM. BIOPHYS. ACTA 1130:81-84(1992).				
CC	-1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL				
CC	CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.				
CC	CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO				
CC	EUKARYOTIC CELL MEMBRANES.				
CC	-1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaborat				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation				
CC	at the European Bioinformatics Institute. There are no restrictions on i				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commerc				
CC	entities requires a license agreement (see http://www.isb-sib.ch/announ				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X60462; G44145; .				
DR	PIR; S22340; S22340.				
DR	PROSITE; PS00481; THIOL_CYTOLYSINS; 1.				
DR	PFAM; PF01289; Thiol_cytolysin; 1.				
DR	HSP; P19995; 1PFO.				
KW	TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.				
FM	SIGNAL	1	25	SEELIGERIOLYSIN.	
FT	CHAIN	26	530	BINDING TO CHOLESTEROL (BY SIMILARITY).	
FT	ACT_SITE	485	485		
SQ	SEQUENCE	530 AA;	59181 MM;	68A497A5 CRC32;	
	Query Match	45.5%;	Score 1502;	DB 1;	Length 530;
	Best Local Similarity	43.1%;	Pred. No. 1.13e-290;		
	Matches 201;	Conservative 115;	Mismatches 149;	Indels 1;	Gaps 1
Dd	62	INKYMCANTDKKSIIVYGQAVTVVPEPKRGKGSGEYIIVEKKKGKGINQNADISVINA	121		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	6	VNDPILAMNDKKKLITLHGSEIEHRFKENQJLPDEFVIEKRKSLSLTNSDIATAT	65		
Dd	122	ISSTIPGALVKANRELVENQPNVLPAKRSLLISVDLPMTIKDKKITVKKPTSSNN	181		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	66	NDSRLYPAGALLVDFETLENNPTLLADRAPMTYSIDLPLASDSDFLOYEDPSMSVVG	125		
Dd	182	AVNTLVGRMWDKSKAYPINAKIDYSDMAYSQSOLIAFGTAFAVANNSLNVFEAIS	241		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	126	AVNDLKMKHMDYGOVN-NVPAPMYEKTITHSMEOALKAFVGSDFEKTSGLDIDFN	184		
Dd	242	DGVQVEVISFKOIIYNVINNEPTSPSKFFGSGVTKEQDLALGMNAENPPATISSVA	301		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	185	SGEROIDIVFRKGIYYTVSAVKNPGDVFDQTVTEDLKORGISAEPRPLAYISSVA	244		
Dd	302	QVYKLISSSHSKYKATAPEAMSKSYKGVVELTNIRKSRKAIYGGSKAKEVEIID	361		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	245	QVYIKLETTSKSOEVAAEAELIKVKAAPTEKKQILDWTGYKAVILGDGSGGARVYT	304		
Dd	362	GNIELDELIDKKSSTDRENPGVPISTYTNLTKONDAVAVANNSEYIETTSKYDGKIN	421		
Oy	:	: : : : : :	: : : : : :	: : : : : :	: : : : : :
Oy	305	GKVMDVEDLDQBSRRTADHPCLPISTYTSFLRNVAATONSNDVYERKVIAYNRGDL	364		
Dd	422	IDHSGYVAOFNISWDEVSYDENGEIKVHKMGCENTKSLAHFTSSIIPLGNARNINITY	481		

```

QY 365 LDHSGAYVAGQYITWNLSDYHOGKEVLTPEKADNRNGODLTAHFTTSIDPLKGVNRLSVK 424
Db 482 ARCTGJFEMWRTVYDRLPLVKNRNVISINCTIYXPHSNVDN 527
QY 425 IRECTGLAMEWRTVYKTDPLVRRRTISIMGTLLYPQVEDKVEN 470

RESULT 6
AC TACY PAEAL STANDARD: PRT: 501 AA.
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE ALVEOLYSIN PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN ALV.
OS PAENIBACILLUS ALVEI (BACILLUS ALVEI).
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
PAENIBACILLUS.
[1]
[1] SEQUENCE FROM N.A.
RX MEDLINE: 91072294.
RA GEOFFROY C., MENGAUD J., ALOUF J.E., COSSART P.;
RT "Alveolysin, the thiol-activated toxin of Bacillus alvei, is
homologous to listeriolysin O, perfringolysin O, pneumolysin, and
streptolysin O and contains a single cysteine."
RL J. BACTERIOL. 172:7301-7305(1990).
[2]
RN SEQUENCE OF 33-60.
RA ALOUF J.E., GEOFFROY C., GILLES A.M., PALMAGNE P.;
RL (IN) RAPPOUILLI R., ALOUF J.E., PALMAGNE P. (EDS.);
NL BACTERIAL PROTEIN TOXINS, PP.49-50, GUSTAV FISHER VERLAG,
STUTTGART (1990).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M62709; GI42473; -.
DR PIR: A37858; A37858.
DR PROSITE: PS00481; THIOL-CYTOLYSINS; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
DR HSSP: P19995; lPFO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 32
FT CHAIN 33 501
FT ACT_SITE 461 461 ALVEOLYSIN.
FT CONFLICT 56 56 BINDING TO CHOLESTEROL (BY SIMILARITY).
SO SEQUENCE 501 AA; 55268 MW; 8DF2C94A CRC32;

Query Match 45.2%; Score 1492; DB 1; Length 501;
Best Local Similarity 42.3%; Pred. No. 1.85e-288;
Matches 192; Conservative 127; Mismatches 132; Indels 3; Gaps 3;

Db 44 IAGLVNRENEVAIGDOIS-FVPEKGIOSNGKFIYVERPKSLTSPVDSITVDSTN 102
QY 10 ILAMVNDKKLLTHGSESTENFT-KEGNOLPDERVIEKRRKSLSTSTISVTAINDS 68
Db 103 RTYPAQLANKDFADNPSLVMAARKPLDISLPGIKNENT-ISOVNPYGYVSSAID 161
QY 69 RLVPALLAVDETLLENPTLLAVDRAPMYSIDPLGLASSDSFLQVEDPSSNSVRAVGN 128
Db 162 QLVSWGKYSSTHTLPRLQVAVESMYSONDISALVNAKVLNGTIGDIFNVAANGCK 221

```

```

QY 129 DLAKWHDVQGVANNVBARQYEKITHASMEQLKVFSGSPFEKTGNSLIDFNSVHSGEK 188
Db 222 KVMVAARKQIFVTVASGLPNNPSDLFPDSDVTFEALARKGVSNAPPLMVSNAVGRITVY 281
QY 189 QIOLVNFKQIYTVYSDAVKNPGDVFQDTVYEDLQKRGISABRPLVYISSVAYGQVYL 248
Db 282 KLETTKSNQVQTAFFKLLNNPSTQASGQYKDIYENSSFAVYVGLGDQCTHNOVYTKDEN 341
QY 249 KLETTSSDVEEAAFEALLIGKVAAPQTEWKQILDNTYKAVILGDPSSGARVYTKVD 308
Db 342 VIOSVIRKDNAQFSSKNPAPYISYTSVFLKDNSTAAVHNNTYEIETETTESKGIKLDHS 401
QY 309 MVELLIQGSFTHDGLPISTTSLRNVATATONSIDYETKVTAYRNGDILLDS 368
Db 402 GAYVAQEVYWDSEFSDADGOEIVTRKSNWDMRDSARHFEPIPLPPAKNIRIFARBC 461
QY 369 GAYVAQYITWNLSDYHOGKEVLTPEKADNRNGODLTAHFTTSIDPLKGVNRLSVIRGC 428
Db 462 TGLAMEWRTVYDEYNPLASDINVSIMGTLLP 495
QY 429 TGLAMEWRTVYKTDPLVRRRTISIMGTLLP 462

RESULT 7
AC TACY STREQ STANDARD: PRT: 574 AA.
ID 054114.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS EQUISIMILIS.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN-SIMD-1;
RX MEDLINE: 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RL TAKAKASHI H., YOTSUDO T.;
RT Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D16824; G498301; -.
DR PROSITE: PS00481; THIOL-CYTOLYSINS; 1.
DR PFAM: PF01289; Thiol_cytolysin; 1.
DR HSSP: P19995; lPFO.
KW TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
FT SIGNAL 1 36
FT CHAIN 37 574
FT ACT_SITE 533 533 STREPTOLYSIN O.
FT CONFLICT 574 574 BINDING TO CHOLESTEROL (BY SIMILARITY).
SO SEQUENCE 574 AA; 63991 MW; 9BBF31B4 CRC32;

Query Match 43.6%; Score 1437; DB 1; Length 574;
Best Local Similarity 42.4%; Pred. No. 2.74e-276;
Matches 194; Conservative 110; Mismatches 151; Indels 3; Gaps 3;

Db 112 INDIKSYLWNELEVLAKNGETIEN-FVPEKGVKRAKRIYVERKKKNINTPVDISIID 170

```

```

OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKRSISTNTSDISVTA 64
DB 171 SYDRTYPALOLANKGFTEKNPDVAVTRKNPKIHIDLPNGDKAT-VEVNDPTYANVS 229
OY 65 TNDRIYPPGALLVDEFTLENNPTLLAVDAPMTYSIDLPGLAASSDFLOYVDPSSSVR 124
DB 230 TAIIDNVOMHNDYSGNTLPARTOYTESMYSKSGIIEALNVNSKILDTGLIDFKSIS 289
OY 125 GAVNDLLAKKHODYGVNNVAPMOKYETITASHMEDLKFKFGSDPEKIGNSLIDIDFNSVH 184
DB 290 KGEKVMIAAYKOIFYTYSANLPNNADYEDKSVTLKEIQRKGVSNENAPPLFVSNVAYGR 349
OY 185 SGEKOIIVNFKOIYTVSVDAVKNPGDVFODTVVEDLKORGISAERPLVYISSVAYGR 244
DB 350 TYVFKLETSSKNDVEAFAEALIKGVDKNTNGKYSIDLENSFTAVVLGGDAEHNKVT 409
OY 245 QVYLKLETSSKNDVEAFAEALIKGVDKNTNGKYSIDLENSFTAVVLGGDAEHNKVT 304
DB 410 KDEPVRNVKIDNATESRKNPAYISTVSFLKNNKIAGVNNREYVETTSYTSKIN 469
OY 305 GKVDMEVDELIOESRFTADHPGLPISTYSFLDNNVATFONSTDVETKVTAYRNGDLL 364
DB 470 LSHOGAVVAOYEILMDEINVDKGEVITKRRDNNWYSKTSFSTVPLGANSRIRIM 529
OY 365 LDHSGAVVAOYITWMLSDHOGKEVLPKANDRNGODLTAFTTSLDKGNVRLSVK 424
DB 530 ARECTGLAMEMWKRVIDERDVKLSEINVISGSLSP 567
OY 425 IRECTGLAMEMWKRVIDERDVKLSEINVISGSLSP 462

RESULT 8
ID TACY_STRPY STANDARD: PRT: 571 AA.
AC P21131;
DT 01-FEB-1991 (REL. 17, CREATED)
DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS PYOGENES.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE; 88057628.
RA KEHOE M.A., MILLER L., WALKER J.A., BOULNOIS G.J.;
RT "Nucleotide sequence of the streptolysin O (SLO) gene: structural
RT homologues between SLO and other membrane-damaging, thiol-activated
RT toxins."
RL INFECT. IMMUN. 55:3228-3232(1987).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18638; G153811; -.
CC PIR; A43507; A43507.
CC PROSITE; PS00481; THIOL_CYTOLYSINS; 1.
CC PFAM; PF01289; Thiol_cytolysin; 1.
CC HSSP; P19995; 1PFO.
CC TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL; PLASMID.
FT SIGNAL 1 33
FT CHAIN 34 571 STREPTOLYSIN O.

```

```

FT ACT SITE 530 530 BINDING TO CHOLESTEROL (BY SIMILARITY).
SQ SEQUENCE 571 AA; 63638 MW; 33124B54 CRC32;
OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKRSISTNTSDISVTA 64
DB 109 INDIKISLVNNELEVLAKNGETIEN-FVREKGVKADKEFVIERKKNINTPPVDISITD 167
OY 6 VNFILAMNYDKKLLTHOGESIENRFI-KEGNOLPDEFVIERKRSISTNTSDISVTA 64
DB 168 SYDRTYPALOLANKGFTEKNPDVAVTRKNPKIHIDLPNGDKAT-VEVNDPTYANVS 226
OY 65 TNDRIYPPGALLVDEFTLENNPTLLAVDAPMTYSIDLPGLAASSDFLOYVDPSSSVR 124
DB 227 TAIIDNVOMHNDYSGNTLPARTOYTESMYSKSGIIEALNVNSKILDTGLIDFKSIS 286
OY 125 GAVNDLLAKKHODYGVNNVAPMOKYETITASHMEDLKFKFGSDPEKIGNSLIDIDFNSVH 184
DB 287 KGEKVMIAAYKOIFYTYSANLPNNADYEDKSVTLKEIQRKGVSNENAPPLFVSNVAYGR 346
OY 185 SGEKOIIVNFKOIYTVSVDAVKNPGDVFODTVVEDLKORGISAERPLVYISSVAYGR 244
DB 347 TYVFKLETSSKNDVEAFAEALIKGVDKNTNGKYSIDLENSFTAVVLGGDAEHNKVT 406
OY 245 QVYLKLETSSKNDVEAFAEALIKGVDKNTNGKYSIDLENSFTAVVLGGDAEHNKVT 304
DB 407 KDEPVRNVKIDNATESRKNPAYISTVSFLKNNKIAGVNNREYVETTSYTSKIN 466
OY 305 GKVDMEVDELIOESRFTADHPGLPISTYSFLDNNVATFONSTDVETKVTAYRNGDLL 364
DB 467 LSHOGAVVAOYEILMDEINVDKGEVITKRRDNNWYSKTSFSTVPLGANSRIRIM 526
OY 365 LDHSGAVVAOYITWMLSDHOGKEVLPKANDRNGODLTAFTTSLDKGNVRLSVK 424
DB 527 ARECTGLAMEMWKRVIDERDVKLSEINVISGSLSP 564
OY 425 IRECTGLAMEMWKRVIDERDVKLSEINVISGSLSP 462

RESULT 9
ID TACY_STRCB STANDARD: PRT: 574 AA.
AC 053957;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE STREPTOLYSIN O PRECURSOR (THIOL-ACTIVATED CYTOLYSIN).
GN SLO.
OS STREPTOCOCCUS CANIS.
OC BACTERIA: FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;
OC STREPTOCOCCUS.
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE; 95102113.
RA OKUMURA K., HARA A., TANAKA T., NICHIGUCHI I., MINAMIDE W.,
RA IGARASHI H., YUTSUDO T.;
RT "Cloning and sequencing the streptolysin O genes of group C and group
RT G streptococci."
RL DNA SEQ. 4:325-328(1994).
CC -1- FUNCTION: SULFHYDRYL-ACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
CC EUKARYOTIC CELL MEMBRANES.
CC -1- SIMILARITY: BELONGS TO THE THIOL-ACTIVATED CYTOLYSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```


DR EMBL: D16825; G498378; -
 DR PROSITE: PS00481; THIOI-CYTOLYSINS; 1.
 DR PFAM: PF01289; Thiol_cytolysin; 1.
 DR HSSP: P19995; 1PRO.
 KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
 FT SIGNAL 1 36 BY SIMILARITY.
 FT CHAIN 37 574 STREPTOLYSIN O.
 FT ACT_SITE 533 533 BINDING TO CHOLESTEROL.
 SO SEQUENCE 574 AA; 63910 MW; 98BD1743 CRC32;
 Query Match
 Best Local Similarity 42.1%; Score 1420; DB 1; Length 574;
 Matches 193; Conservative 108; Mismatches 154; Indels 3; Gaps 3;
 Db 112 INKITYSLNNEVLAKNETTEN-FVPEKGYKAKDKTIVIERKKNNITPVDISTID 170
 QY 6 VNDPILAMNYDKKKLLTHOGESEIENRFI-KEGNQLPDEFVIERKKRSISTSTDSISVA 64
 Db 171 SVTDRTYPALQLANKGTEENKPDVAVTRKPKIHIDPGMGDKAT-VEVNDPTAYNS 229
 QY 65 TNSRLYPGALLVDETLLENPTLLAVRAPATYSIDLPGLASSDSFLOVEDPSNSVVR 124
 Db 230 TAIIDNVNOMHNDYSGCNTLPARTQYTESNYSKQIEALNVNSKILDTGLIDFKSIS 289
 QY 125 GAVNDLLAKWHODYGVNNVPAHQYEKITAHSMEQLKVFSGDEFKTSNIDIDENSVAH 184
 Db 290 KGEKYMIAAYKQIEFTVSANLPNNPDVFDKSVTFKELQAGVSNAPPLVSNVAVGR 349
 QY 185 SGKQIOIVNFQOITYTSVDAYKNGDVFODTVYEDLKORISERPLVYISSVAVGR 244
 Db 350 TVFKLETSKSDVBAASALKGFDVTKNGKSDILENSFTAVLGADAENHKVYT 409
 QY 245 QYALKETTSKSDVBAEALIKGVKAPQTEEMKOILDTEVKAVILGDDSSGARVYT 304
 Db 410 KPDVIRNVIKANATSKRNPAPISTYSVFLKNNIAGVNNRSEVETSTEYTSKIN 469
 QY 305 GKVDWEDDIOESRFTADHPGLPISTYSFLDNVAFQNSTDVEETKVAIRNGDGL 364
 Db 470 LSHQGVVAVOYELTMEINVDKGEVITKRRDNWYKTSFESVITPLGANSRIRIM 529
 QY 365 LSHQGVVAVOYELTMEINVDKGEVITKRRDNWYKTSFESVITPLGANSRIRIM 529
 Db 530 ARECTGLAEMWRTVDEKVLKSEININISSTLSP 567
 QY 425 IRECTGLAEMWRTVDEKVLKSEININISSTLSP 567
 RESULT 10
 ID TACY_BACCE STANDARD; PRT; 485 AA.
 AC Q45105;
 DT 15-JUL-1998 (REL. 36, CREATED)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 CC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;
 CC BACILLUS.
 CC SEQUENCE FROM N.A.
 CC STRAIN-RIMD 206001;
 CC YUTSUDO T.;
 CC SUBMITTED (NOV-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: SUFFHYDRACTIVATED TOXIN. IS ABLE TO LYSE CHOLESTEROL
 CC CONTAINING MEMBRANES. CAN BE REVERSIBLY INACTIVATED BY OXIDATION.
 CC CHOLESTEROL IS THE RECEPTOR FOR THE BINDING OF THESE TOXINS TO
 CC EUKARYOTIC CELL MEMBRANES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE THIOI-ACTIVATED CYTOLYSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D21270; G600252; -
 DR PROSITE: PS00481; THIOI-CYTOLYSINS; 1.
 DR PFAM: PF01289; Thiol_cytolysin; 1.
 DR HSSP: P19995; 1PRO.
 KM TOXIN; HEMOLYSIS; CYTOLYSIS; SIGNAL.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 >485 HEMOLYSIN.
 FT ACT_SITE 466 466 BINDING TO CHOLESTEROL (BY SIMILARITY).
 FT NON_TER 485 485
 SO SEQUENCE 485 AA; 53862 MW; 45208ABA CRC32;
 Query Match
 Best Local Similarity 40.1%; Score 1358; DB 1; Length 485;
 Matches 175; Conservative 121; Mismatches 137; Indels 3; Gaps 3;
 Db 44 SIDGIGNLTYNNOEVLAVNGDKYES-FVPEKESINSNGKRVVVDVRKHLQRHOSIFRL 102
 QY 5 AVNDFILAMNYDKKKLLTHOGESEIENRFI-KEGNQLPDEFVIERKKRSISTSTDSISVA 63
 Db 103 DSVANRTYPGAVOLANKAPADNOPSLVAKRKPLNTSIDPGRKENT-ITVQNPYGVN 161
 QY 64 ATNDSRLYPGALLVDETLLENPTLLAVRAPATYSIDLPGLASSDSFLOVEDPSNSVVR 123
 Db 162 AGAVDDLVSTWNEKYSATHTLPARMQYTESMYSKQAIASALNVNKKYLDNSINIPNAV 221
 QY 124 RGAVNDLLAKWHODYGVNNVPAHQYEKITAHSMEQLKVFSGDEFKTSNIDIDENSVAH 183
 Db 222 ANGKRYMIAAYKQIEFTVSANLPNNPDVFDKSVTFKELQAGVSNAPPLVSNVAVGR 281
 QY 184 HSGKQIOIVNFQOITYTSVDAYKNGDVFODTVYEDLKORISERPLVYISSVAVGR 243
 Db 282 RTVYFKLETSKSDVBAASALKGFDVTKNGKSDILENSFTAVLGADAENHKVYT 341
 QY 244 ROYALKETTSKSDVBAEALIKGVKAPQTEEMKOILDTEVKAVILGDDSSGARVYT 303
 Db 342 TKDENEIRNIITKDAELSFKNPAPISTYSFLKDNATAVHNNTYIETTESYSAK 401
 QY 304 TGKVDWEDDIOESRFTADHPGLPISTYSFLDNVAFQNSTDVEETKVAIRNGDGL 363
 Db 402 TLHIGAVVAVOYELTMEINVDKGEVITKRRDNWYKTSFESVITPLGANSRIRIM 461
 QY 364 LLDHSGAVVAVOYELTMEINVDKGEVITKRRDNWYKTSFESVITPLGANSRIRIM 461
 Db 462 VARECTGLAEMWRTV 477
 QY 424 KIRECTGLAEMWRTV 439
 RESULT 11
 ID AOPA_BOVIN STANDARD; PRT; 527 AA.
 AC P21398;
 DT 01-MAY-1991 (REL. 18, CREATED)
 DT 01-MAY-1991 (REL. 18, LAST SEQUENCE UPDATE)
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
 DE AMINE OXIDASE [FLAVIN-CONTAINING] A (EC 1.4.3.4) (MONOAMINE OXIDASE)
 GN MAOA.
 CC BOS TAURUS (BOVINE).
 CC EDAROTIA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 CC ARTIODACTYLA; RUMINANTIA; PECORA; BOVOIDEA; BOVIDAE; BOVINAE; BOS.
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE: 89246344.
 CC POWELL J.F., HSU Y.P., WEYLER W., CHEN S., SALACH J.,
 CC ANDRIKOPOULOS K., MALLETT J., BREAKFIELD X.O.;
 CC "The primary structure of bovine monamine oxidase type A. Comparison
 CC with peptide sequences of bovine monamine oxidase type B and other
 CC flavoenzymes.";
 CC RT BIOCHEM. J. 259:407-413(1989).


```

CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEAMINATION OF BIOGENIC AND
CC XENOBIOTIC AMINES AND HAS IMPORTANT FUNCTIONS IN THE METABOLISM OF
CC NEUROACTIVE AND VASOACTIVE AMINES IN THE CENTRAL NERVOUS SYSTEM
CC AND PERIPHERAL TISSUES. MAO-A PREFERENTIALLY OXIDIZES BIOGENIC
CC AMINES SUCH AS 5-HYDROXYTRYPTAMINE (5-HT), NOREPINEPHRINE AND
CC EPINEPHRINE.
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: MONOMER. HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMATICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15609; G525; ALT_INIT.
DR PIR: S03974; S03974.
KW OXIDOREDUCTASE; FLAVOPROTEIN; FAD; TRANSMEMBRANE; MITOCHONDRION;
KW NEUROTRANSMITTER DEGRADATION.
FT NP_BIND 14 70 FAD (ADP PART) (POTENTIAL).
FT BINDING 406 406 FAD (BY SIMILARITY).
FT TRANSMEM 498 518 POTENTIAL.
SQ SEQUENCE 527 AA; 59800 MW; 724E4396 CRC32;

Query Match
Best Local Similarity 20.0%; Score 124; DB 1; Length 527;
Matches 30; Conservative 42; Mismatches 72; Indels 6; Gaps 5;

Db 1 MESLQTSDAQGFVYVYIGGIGSLAAK-LLAHEVAVVLEAREYRGVGGTYVNRH 59
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 LETTSKDEVEAFALINGVKAAPQTEWKOILDNTEVKAIVLGGPSSGAVVTKYDM 309
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 V-DYVVGAGVYVPTQNRILRLSKQLEFYKVNVERLVHVYKGTPEFARPPVWNP 118
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 310 VEDLIDEGSFTFDHGLPLSYTTSF-LPDNVYATFONSTDEYETKVTAYRKGDLIDHS 368
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 IAVL-DYNNLMRTM--DNMGKEIPADAPWE 145
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 GAVVAQYVITWNLSTYDHQKEVLTPKAMD 398
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
ID NOPS_YEAST STANDARD; PRT; 511 AA.
AC 012499;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
FT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
FE NUCLEOLAR PROTEIN NOPS.
GN NOPS OR YOR310C OR O6108.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOCMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=9288C / FY1679;
RA PEARSON B.M., HERNANDO Y., WOLF S.S., KALOGEROPOULOS A., SCHWEIZER M.;
RT SUBMITTED (AUG-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC [2]
CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
CC STRAIN=M303;
CC MEDLINE: 98298165.
CC WU P., BROCKENBROUGH J.S., METCALFE A.C., CHEN S., ARIS J.P.;
CC "Nops is a small nuclear ribonucleoprotein component required for
CC pre-18S rRNA processing in yeast."

```

```

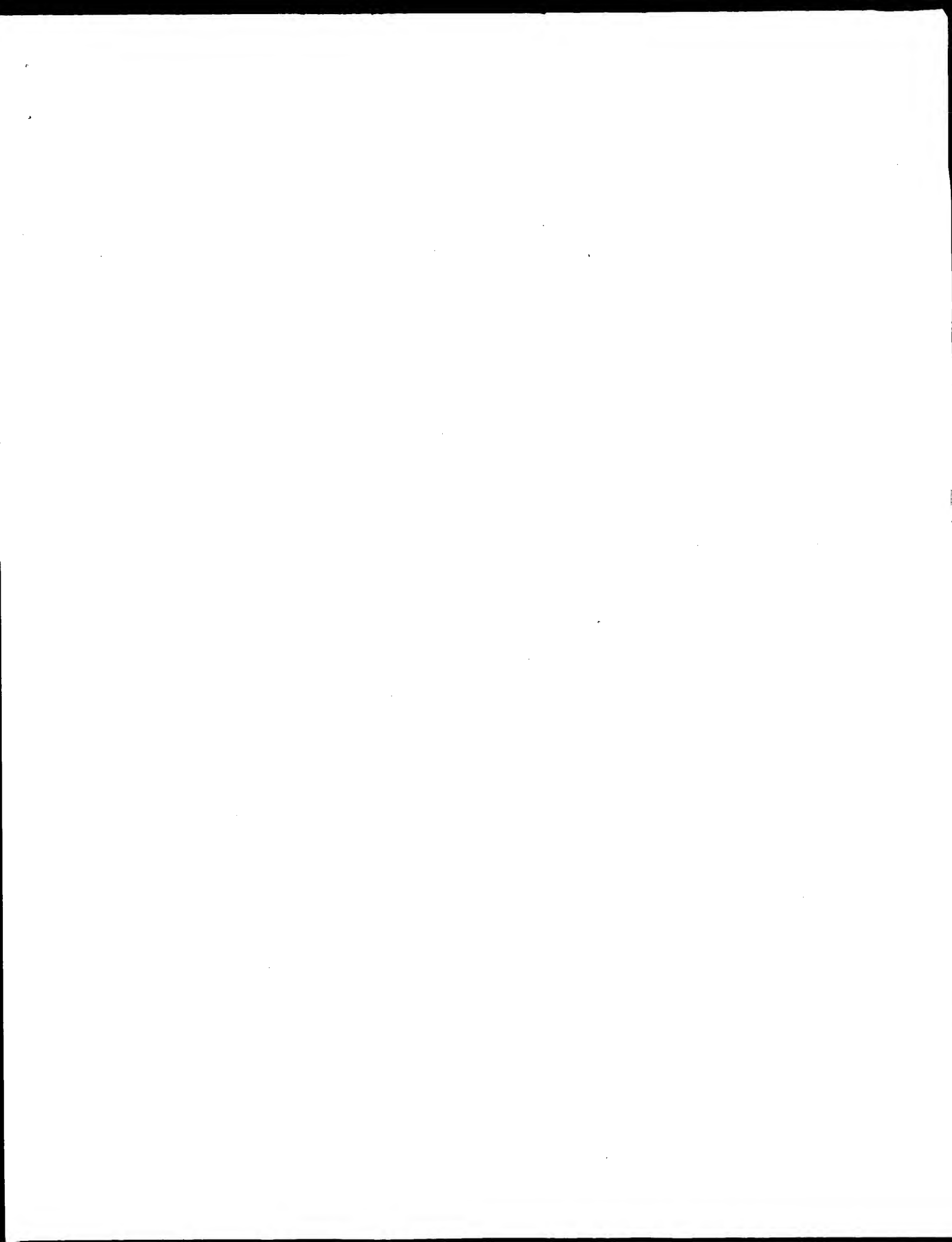
RL J. BIOL. CHEM. 273:16453-16463(1998).
CC -1- FUNCTION: REQUIRED FOR PRE-18S RNA PROCESSING.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: BELONGS TO THE NOPS/NOPS6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X90565; G940841; -.
DR EMBL: 275217; E252143; -.
DR SGD: AF056070; G3170533; -.
DR NUCLEAR PROTEIN; RNA PROCESSING.
FT DOMAIN 441 511 ASP/GLU/LYS-RICH.
SQ SEQUENCE 511 AA; 56956 MW; F04AF72C CRC32;

Query Match
Best Local Similarity 22.8%; Score 112; DB 1; Length 511;
Matches 31; Conservative 39; Mismatches 58; Indels 8; Gaps 8;

Db 188 YGMH-FPEELAKIV-TDSVAVARILLMGIRSKASETDLSEILPEELER-VYTAAEVSMG 244
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 YGROYVLEKLTETTSKDEVEAFALINGVKA-APQTEWKOILDNTEVKAIVLGGPSSG 299
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 TETITQDLDNINLARGVYEFARVREOLS-NYLSAKMKAIAPLQVLQELGARLIA-H 302
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 300 ARVYTGAVDVEDLIDGSGFTADHGLPLSYTTSFLPDNVYATFONSTDEYETKVTAYR 359
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 SGLSLIASKSPASTIQ 318
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 360 NGDLL-LDHS GAVVAQ 374
: : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
ID DE3A_MYCPN STANDARD; PRT; 872 AA.
AC P75404;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE DNA POLYMERASE III, ALPHA CHAIN (EC 2.7.7.7).
DN DNAE.
OS MYCOPLASMA PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RS SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE: 97105885.
RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA HERRMANN R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC -1- FUNCTION: DNA POLYMERASE III IS A COMPLEX, MULTICHAIN ENZYME
CC RESPONSIBLE FOR MOST OF THE REPLICATIVE SYNTHESIS IN BACTERIA.
CC THIS DNA POLYMERASE ALSO EXHIBITS 3' TO 5' EXONUCLEASE ACTIVITY.
CC THE ALPHA CHAIN IS THE DNA POLYMERASE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =
CC N PYROPHOSPHATE + DNA(N).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

WISKEY

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_bp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Aug 30 16:05:23 1999; MasPar time 29.10 Seconds

Tabular output not generated. 883.456 Million cell updates/sec

Title: >US-09-120-04-3

Description: (1-471) from US09120044.pep

Perfect Score: 3299

Sequence: 1 NANKAVNDILAMNYDKKKL.....TISWGTITLPQVEDKVEN

Scoring table: PAM 150

Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: sptrembl9

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 49.610; Variance 95.041; scale 0.522

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1808	54.8	497	2	055996	SULIYSIN.	0.00e+00
2	1803	54.7	497	2	085102	HEMOLYSIN.	0.00e+00
3	1329	40.3	534	2	031241	POLYSIN.	4.77e-251
4	124	3.8	50	2	048772	LISTERIOLYSIN O (FRAGM	2.46e-03
5	118	3.6	50	2	085906	LISTERIOLYSIN O (FRAGM	1.69e-02
6	118	3.6	1361	2	030524	S-LAYER PROTEIN.	1.69e-02
7	118	3.6	1361	2	087083	S-LAYER PROTEIN.	1.69e-02
8	117	3.5	265	2	012044	CHROMOSOME XV READING	2.32e-02
9	114	3.5	450	2	068518	HYPOHETICAL 49.7 KD P	1.69e-02
10	111	3.4	831	3	013705	HYPOHETICAL 95.0 KD P	1.48e-01
11	109	3.3	744	2	067108	DNA GYRASE A SUBUNIT.	2.71e-01
12	110	3.3	867	10	081747	HYPOHETICAL 99.2 KD P	2.01e-01
13	106	3.2	187	1	058072	187AA LONG HYPOHETICA	6.62e-01
14	104	3.2	231	11	035843	KERATIN 14 (FRAGMENT).	1.19e+00
15	104	3.2	303	3	005541	CHROMOSOME IV COSMID 9	1.19e+00
16	106	3.2	501	1	028947	CONSERVED HYPOHETICAL	6.62e-01
17	107	3.2	710	2	085906	TRANSFERRIN PROTEIN B.	4.93e-01
18	106	3.1	1029	10	080955	T16B24.10 PROTEIN.	6.62e-01
19	102	3.1	142	2	085634	L0038.	2.11e+00
20	102	3.1	277	2	051546	PHOSPHATE TRANSPORT AT	2.11e+00

21	101	3.1	362	2	050434	AMINOTRANSFERASE.	2.80e+00
22	102	3.1	446	2	034363	PUTATIVE L-AMINO ACID	2.11e+00
23	102	3.1	1329	5	076356	C45G7.6 PROTEIN.	2.11e+00
24	101	3.1	5027	11	063460	PROLINE-RICH PROTEIN (2.80e+00
25	100	3.0	124	5	017618	C04B4.6 PROTEIN.	3.72e+00
26	100	3.0	142	2	052137	ROPER.	3.72e+00
27	98	3.0	165	13	P87487	GROWTH HORMONE (FRAGME	6.50e+00
28	99	3.0	218	1	038691	HYPOHETICAL PROTEIN M	4.92e+00
29	99	3.0	259	1	027868	CELL DIVISION INHIBITO	4.92e+00
30	99	3.0	495	2	033330	HYPOHETICAL 53.3 KD P	4.92e+00
31	98	3.0	508	5	025917	MITOGEN-ACTIVATED PROT	6.50e+00
32	98	3.0	539	10	080914	MYOSIN HEAVY CHAIN-LIK	6.50e+00
33	100	3.0	640	5	009647	HYPOHETICAL 74.1 KD P	6.50e+00
34	98	3.0	688	5	045430	F32B6.2 PROTEIN.	6.50e+00
35	100	3.0	814	10	065715	HYPOHETICAL 91.9 KD P	6.50e+00
36	98	3.0	865	2	P73260	HYPOHETICAL 97.4 KD P	6.50e+00
37	100	3.0	1035	2	025887	CATION EFFLUX SYSTEM P	3.72e+00
38	99	3.0	1359	10	081726	HYPOHETICAL 155.3 KD	4.92e+00
39	99	3.0	1817	5	019931	COSMID F31D5.	4.92e+00
40	97	2.9	315	2	P70803	RESTRICTION ENDONUCLEA	8.56e+00
41	97	2.9	344	2	Q48027	OUTER MEMBRANE PROTEIN	8.56e+00
42	97	2.9	358	2	P74134	AT103.	8.56e+00
43	96	2.9	437	2	P74770	7-KETO-8-AMINOPELAGON	1.12e+01
44	96	2.9	1123	14	069089	UL37.	1.12e+01
45	97	2.9	2160	5	017709	H05L14.2 PROTEIN.	8.56e+00

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	497 AA.
AC	055996	055997:		
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST SEQUENCE UPDATE)		
DT	01-JAN-1999	(TREMBLREL. 09, LAST ANNOTATION UPDATE)		
DE	SULIYSIN.			
GN	SLY.			
OS	STREPTOCOCCUS SUTS.			
OC	BACTERIA, FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;			
OC	STREPTOCOCCUS.			
RF	(1)			
RP	SEQUENCE OF 1-71 FROM N.A.			
RC	STRAIN-P1/7:			
RA	SEGERS R.P.A.M., KENTER T., DE HAAN W.A.M., JACOBS A.A.C.;			
RT	"Characterization of the gene encoding suliyisin from Streptococcus			
RT	suts, and expression in field strains."			
RL	SUBMITTED (OCT-1998) TO EMBL/GENBANK/DBJ DATA BANKS.			
DR	EMBL; Z36907; E1334304; -			
SQ	SEQUENCE 497 AA; 54850 MW; D78AEBEF CRC32;			

Query Match	54.8%; Score 1808; DB 2; Length 497;
Best Local Similarity	50.4%; Pred. No. 0.00e+00;
Matches	236; Conservative 113; Mismatches 117; Indels 2; Gaps 2;
Db	29 SKODINOFOSLYTEPQELTNEGEYIDNPATGMLENGFRVYLREKKNITNNSADIA 88
QY	2 NANKAVNDILAMNYDKKKLTHGSELENRPKEGNLPDEFVIERKKSLSTNTSDIS 61
Db	89 VIDAKAANIYPGALLRADONLDNNPTLISARGDLTSLNPLANGDSHTVANSPTFS 148
QY	62 VTATNDSRLYPGALLVYDELTENPLLVADRAPMTYSIDLPLGASSDSFLOVEDPSNS 121
Db	149 TVRQGVNMLSKRNNTYAGEGNTGALOEYDETMAYMSOLTKFTGTSFEKIAVPLDINF 208
QY	122 SVRAQVNDLAKRWHDY-GOVNVPARMOYEKLTFAHSMELKVFSGDFEKTGNSIDIDF 180
Db	209 DAVNSGEKOVIVNEKOIYTVSVDEPESPSKLEASTTJEDLKRKNITTEVPYVSVY 268
QY	181 NSVHSGEKQIDIVAFKQIYTVSDAVKNGVDVFDVTVEDLKKRQISERPLVYISVY 240
Db	269 SYGSMFKLETSSRSSTOVQAAFKAKGVDSIGMAEYODILKNTSSSAIFGSDGSA 328
QY	241 AYGRQVYLKLETTSKSDSEVAEFAELIKGVKAPQTEWKRQILNTEKVAIILGDISSGA 300

Query	Best Local Similarity	Conservative	Score	DB 2	Length	534	
Query Match	40.3%	125	1329	DB 2	Length	534	
Best Local Similarity	40.7%	125	1329	DB 2	Length	534	
Matches	192	Conservative	125	Mismatches	147	Indels	8
							Gaps
							6
Db	63	ENGVCATKRGKLYKYPSSGLAVKGSIEENVPTK-DOLKDGTYTFKHERSENNLRSDIS	121				
Qy	3	NKAVNDFLMAYDKKLLTHQGESIENRFIRKGNLPDE-FVIERKKSLSTINTDIS	61				
Db	122	AFDANNAHVPGALVLANKDLAKGSPISIGARAPQTVSVDLPGLVDGKNKRVINNPFS	181				
Qy	62	VYATNDSRLYPCALLVDETLENNPRLTADVRAPMTYSIDLPLGASDSDFLQVEDPSNS	121				
Db	182	SYTGQNLGDLGDMTORNKSPYDPHAKISYDETMTSKROLKAGLGEKYSAKLVNDF	241				
Qy	122	SVRGAVNDLAKWHDQYQVNNVPARMQYERTITASHMQLVKYKGSDFEKGNSLIDFN	181				
Db	242	AIHKKEEROVALISFQIYYTASVDPPTSPHSVFGPNVTAODLKGYNKKNPLGYISSVS	301				
Qy	182	SVHSEKQOQIVNFQOIYYTVSDAVKKNPGVQFDYTVEDLKRGISAEERPLYISSVA	241				
Db	302	YGRQIFVKLETTSTNDYQAAFSGLFKAFKGNLSIEFAKADILNKTRATVAVGSGAR	361				
Qy	242	YGRQYVLYKLETTSEKDEVEAEALIKG--VKVAPQ--TEKKQILDNEVKAVIDGDP	297				
Db	362	GGVEVATGIDMLKTKIEESTSYKRVAVPVSQVNFKNLQNLAAVSSGSDYIETTAIT	421				
Qy	298	SGARVATGVDVDEDLIGESRFTADHGRPLSYTSLRDNVAVATFQNSTDYETKVTA	357				
Db	422	YKSGELTFHGSGYVAKFRLKWDEISYDPQCKEIRPTKWSGNNAARLGGREITQLPAN	481				
Qy	358	YKSGELTFHGSGYVAKFRLKWDEISYDPQCKEIRPTKWSGNNAARLGGREITQLPAN	417				
Db	482	ARNIHVEAGATGLAMPWM-TVINKKMLPLVPHBEIYLGKGTLLPWPEDNV	532				
Qy	418	VRNLVSKIRECTGLAME-WKRTVYKTDLPVYKRTISTWGTLLPYQVEDKY	468				
RESULT	4	PRELIMINARY	PRT	50	AA		
ID	048772						
AC	048772						
DT	01-NOV-1996	(TREMBLREL. 01, CREATED)					
DT	01-NOV-1996	(TREMBLREL. 01, LAST SEQUENCE UPDATE)					
DE	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)					
GN		HLV					
OS		LISTERIA MONOCYTOGENES.					
OC		BACTERIA: FIRMICUTES: BACILLUS/CLOSTRIDIUM GROUP; BACILLACEAE;					
OC		LISTERIA.					
RP		[1]					
RN		SEQUENCE FROM N.A.					
RC		STRAIN-11984, TYPE 1;					
RX		MEDLINE; 9611685.					
RA		RASMUSSEN O.F., SKOUBOE P., DONS L., ROSSEN L., OLSEN J.E.;					

[illegible]


```

RC SEQUENCE FROM N.A.
RA STRAIN-S288C (AB972);
RL FULTON L.;
RN SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

[2]
RN SEQUENCE FROM N.A.
RA STRAIN-S288C (AB972);
RL JONATHAN M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DU Z.,
RN FAVELLO A., FULTON L., GATUNG S., GRECO T., KIRSTEN J., KUCABA T.,
RA HARTSMORTH K., HARKINS J., HILLER L., JIER M., JOHNSON D.,
RN JOHNSTON L., LANGSTON Y., LATREILLE P., LE T., MARDIS E., MENEZES S.,
RA MITLER N., NHAN M., PAULEY A., PELUSO D., RIEKEN L., RILES L.,
RN TAICH A., TREVASKIS E., VIGNATI D., WILCOX L., WOHLDMAN P., VAUDIN M.,
RA WILSON R., WATERSTON R.;
RN SUBMITTED (MAR-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

[3]
RN SEQUENCE FROM N.A.
RA STRAIN-S288C (AB972);
RL WATERSTON R.;
RN SUBMITTED (MAY-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

[4]
RN SEQUENCE FROM N.A.
RA STRAIN-S288C (AB972);
RL JIA Y., CHERRY J.M.;
RN SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 051031; G1332636; -
SQ SEQUENCE 303 AA; 33964 MW; D494B483 CRC32;

```

```

Query Match 3.2%; Score 104; DB 3; Length 303;
Best Local Similarity 29.9%; Pred. No. 1.19e+00;
Matches 26; Conservative 26; Mismatches 27; Indels 8; Gaps 5;

DB 84 YGFELQGLPSKNNMAGNGSNSNTNKSMPEDLGHRAQKFIILNNVPHSKNFDDEKILQS 143
QY 99 YSIDLPGLASDSDFLEVEDPSNSVGVAVNDLAKWHDYGGVNNVPARMOYE--KI--T 154
DB 144 AHTEELIVGTETIGDDIASGTSNLE 170
QY 155 AHSMEQLKVK--F-GSDPEK-TGNSLD 177

```

Search completed: Mon Aug 30 16:06:48 1999
Job time : 85 secs.